

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 1999, 11:20:12 ; Search time 13.2 Seconds
(without alignments)
16.150 Million cell updates/sec

Title: US-08-819-669D-26

Perfect score: 52

Sequence: 1 EADPTGHSY 9

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	9	1 R29769	Antigen E peptide.
2	52	100.0	9	1 R49224	HLA-A1 MAGE 1 anti
3	52	100.0	9	1 R47330	HLA-A1 MAGE 1 anti
4	52	100.0	9	1 R50281	MAGE-1 nonapeptide
5	52	100.0	9	1 R36675	Synthetic peptide
6	52	100.0	9	1 R5135	MAGE 1 immunogenic
7	52	100.0	309	1 R70909	Human melanoma ant
8	52	100.0	9	1 R65112	MAGE 1 immunogenic
9	52	100.0	9	1 R75954	Melanoma antigen (
10	52	100.0	9	1 R82988	P815 antigenic pep
11	52	100.0	9	1 R78824	MAGE-1 cytotoxic T
12	52	100.0	12	1 R80620	Immunogenic peptid
13	52	100.0	9	1 R83932	MHC class I restri
14	52	100.0	9	1 R90692	Human leukocyte an
15	52	100.0	9	1 R93343	MAGE-1 nonapeptide
16	52	100.0	9	1 W00897	Human melanoma MAG
17	52	100.0	10	1 W23038	MAGE-1/HLA-B44 tum
18	52	100.0	9	1 W36729	MAGE-1 antigenic p
19	52	100.0	9	1 W54622	Peptide from MAGE-
20	52	100.0	9	1 W68371	Human MAGE-1 pepti
21	52	100.0	9	1 W78838	MAGE-1 protein fra
22	52	100.0	9	1 W77125	gp75/TRP-1 synthet
23	52	100.0	9	1 W57336	Peptidase-resistan
24	52	100.0	9	1 W57336	Peptidase-resistan
25	52	100.0	309	1 W81548	Tumour rejection a
26	52	100.0	9	1 W89945	HLA-A1 binding pep
27	52	100.0	9	1 Y00685	Tumour antigen boo
28	52	100.0	9	1 Y10633	Peptide antigen SE
29	52	100.0	9	1 Y10623	Peptide antigen SE
30	52	100.0	9	1 Y10424	HLA Class I motif
31	52	100.0	9	1 Y01727	Exemplary antigeni
32	49	94.2	9	1 R93342	HLA binding nonape
33	48	92.3	9	1 W57333	Peptidase-resistan
34	48	92.3	9	1 W57335	Peptidase-resistan
35	47	90.4	9	1 R99337	HLA binding nonape
36	47	90.4	9	1 R99339	HLA binding nonape
37	46	88.5	9	1 R99340	HLA binding nonape
38	44	84.6	9	1 R99338	HLA binding nonape
39	42	80.8	9	1 R99341	HLA binding nonape
40	42	80.8	9	1 Y10628	Peptide antigen SE
41	41.5	79.8	10	1 W54604	Peptide 1 from MAg
42	40	76.9	9	1 Y10629	Peptide antigen SE
43	37	71.2	88	1 W16332	Baboon MAGE-3 homo

ALIGNMENTS

```
RESULT 1
R29769
ID R29769 standard; Peptide; 9 AA.
AC R29769:
DT 22-APR-1993 (first entry)
DE Antigen E peptide.
KW Antigen; tumorigenic cell; A+ B+; T-cell; response; syngeneic;
KW animal; mouse; tumour rejection antigen precursor; TRAP; PIA.
OS Homo sapiens.
PN W09220356-A.
PD 26-NOV-1992.
PF 22-MAY-1992; 004354.
PR 23-MAY-1991; US-705702.
PR 09-JUL-1991; US-728838.
PR 23-SEP-1991; US-764364.
PR 12-DEC-1991; US-807043.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;
PI Van Den Eynde B, Van Der Bruggen P, Van Pel A;
DR WPI; 92-415460/50.
PT Nucleic acid mol. encoding a human tumour rejection antigen
PT precursor - useful as an immunostimulant in a vaccine for
PT treating and preventing cancers, also useful in diagnosis
PT Disclosure: Page 97; 142pp; English.
PS This sequence represents the sequence of the antigen E. Antigens such
CC as this one cause a T-cell response to be elicited which transplanted
CC into a syngeneic animal, usually a mouse. This antigen is derived from
CC the cell line MEL-3.1. See also Q32351.
CC Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
Db 1 EADPTGHSY 9

RESULT 2
R49224
ID R49224 standard; Protein; 9 AA.
AC R49224:
DT 31-AUG-1994 (first entry)
DE HLA-A1 MAGE 1 antigen peptide fragment 958.01.
KW Immunogenic; HLA-A3.2; HLA-A1; binding motif; MHC molecule;
KW immune response; viral infection; cancer; prostate cancer; lymphoma;
KW hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS Synthetic.
PN W09403205-A.
PD 17-FEB-1994.
PF 06-AUG-1993; 007421.
PR 07-AUG-1992; US-926666.
PR 05-MAR-1993; US-027746.
PA (CYTE-) CYTEL CORP.
PI Celis E, Grey HM, Kubo RT, Sette A;
PI WPI; 94-065403/08.
DR Peptide which specifically binds selected MHC allele - used to
PT induce an immune response for treatment or prevention of viral
PT infection or cancer, or for diagnosis.
PS Example 16; Page 116; 150pp; English.
CC The sequences given in R47304-33 and R49201-44 are immunogenic
CC peptides which have a HLA-A3.2, HLA-A1 or a HLA-A1 binding motif.
CC These peptides may be used in the composition of the invention.
CC These peptides are capable of binding selected MHC molecules and
```

44 37 71.2 91 1 W16333 Baboon MAGE-3 homo
45 37 71.2 126 1 W16334 Baboon MAGE-3 homo

CC inducing an immune response. They can be used to treat and/or
 CC prevent viral infection and cancer, eg. prostate cancer, lymphoma,
 CC hepatitis or AIDS. They can also be used to produce antibodies for
 CC use as diagnostic or therapeutic agents. The peptides can also be
 CC used as diagnostic agents.
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 | | | | |
 Db 1 EADPTGHSY 9

RESULT 3
 R47330 ID R47330 standard; Protein; 9 AA.
 AC R47330;
 DE 31-AUG-1994 (first entry)
 KW HLA-A1 MAGE 1 antigen peptide fragment 161-169.
 KW immunogenic; HLA-A3.2; HLA-A1; binding motif; MHC molecule;
 KW immune response; viral infection; cancer; prostate cancer; lymphoma;
 KW hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
 OS Synthetic.
 PN W09403205-A.
 PD 17-FEB-1994.
 PF 06-AUG-1993; U07421.
 PR 07-AUG-1992; US-928666.
 PR 05-MAR-1993; US-027746.
 PA (CYTE-) CYTEL CORP.
 PI Cellis E, Grey HM, Kubo RT, Sette A;
 DR WPI: 94-065403/08.

PT Peptide which specifically binds selected MHC allele - used to
 PT induce an immune response for treatment or prevention of viral
 PT infection or cancer, or for diagnosis
 PS Example 8; Page 52; 150pp; English.
 CC The sequences given in R47304-33 and R49201-44 are immunogenic
 CC peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
 CC These peptides may be used in the composition of the invention.
 CC These peptides are capable of binding selected MHC molecules and
 CC inducing an immune response. They can be used to treat and/or
 CC prevent viral infection and cancer, eg. prostate cancer, lymphoma,
 CC hepatitis or AIDS. They can also be used to produce antibodies for
 CC use as diagnostic or therapeutic agents. The peptides can also be
 CC used as diagnostic agents.
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 | | | | |
 Db 1 EADPTGHSY 9

RESULT 4
 R50281 ID R50281 standard; Protein; 9 AA.
 AC R50281;
 DE 26-SEP-1994 (first entry)
 KW MAGE-1 nonapeptide.
 KW MAGE; nonapeptide; cancer; melanoma; breast cancer; HLA;
 KW histocompatibility; human leucocyte antigen; probe; treatment;
 KW therapy; vaccine.
 OS Synthetic.
 PN W09403304-A.
 PD 17-MAR-1994.
 PF 30-AUG-1993; U08157.

PR 31-AUG-1992; US-938334.
 PR 26-MAR-1993; US-037230.
 PR 07-JUN-1993; US-073103.
 PA (LUDW-) LUDWIG INST CANCER RES.
 PI Boon-falleur T, De Plaen E, Lurquin C, Traversari C;
 PI Van Derbruggen P;
 DR WPI: 94-100844/12.
 DR N-PSDB: Q44751.

PT New nona-peptide derived from tumour rejection antigen precursor
 PT - presented by HLA-A1 cancer cells, for use in diagnosis or
 PT therapy of esp. melanoma and breast cancer.
 PS Disclosure; Page 19; 33pp; English.
 CC An isolated nonapeptide having the amino acid sequence Glu-Val-Asp-
 CC Pro-Ile-Gly-His-Leu-Tyr is derived from the tumour rejection antigen
 CC precursor encoded by the MAGE-3 gene and presented by HLA-A1. The
 CC nonapeptide can be used in a vaccine to treat a cancerous condition
 CC involving HLA-A1 subtype cancerous cells. The nucleic acid encoding
 CC the nonapeptide can be used as a probe to identify tumour cells.
 CC This sequence is homologous to the peptide described and is encoded
 CC by the MAGE-1 gene.
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 | | | | |
 Db 1 EADPTGHSY 9

RESULT 5
 R63675 ID R63675 standard; Protein; 9 AA.
 AC R63675;
 DE 22-JUN-1995 (first entry)
 DE Synthetic peptide derived from exon 3.1 of MAGE 1.
 KW Melanoma antigen-1; MAGE-1; cytolytic T cells; antigen E; exon 3.1.
 OS Synthetic.
 PN W09423031-A.
 PD 13-OCT-1994.
 PF 17-MAR-1994; U02877.
 PR 26-MAR-1993; US-037230.
 PA (LUDW-) LUDWIG INST CANCER RES.
 PI Boon-falleur T, Gaugler B, Van DEN EYNDE B, Van DER BRUGEN P;
 DR WPI: 94-333192/41.
 PT New tumour rejection antigen precursor MAGE3 - useful in
 PT treatment and diagnosis of cancer
 PS Example 34; Page 36; 105pp; English.

CC R63675 is a synthetic peptide derived from exon 3.1 of melanoma
 CC antigen-1 (MAGE-1), it was used to transfer antigen-E cytolytic T
 CC lymphocyte sensitivity to normally non-sensitive cells.
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 | | | | |
 Db 1 EADPTGHSY 9

RESULT 6
 R65135 ID R65135 standard; peptide; 9 AA.
 AC R65135;
 DE 09-OCT-1995 (first entry)
 DE MAGE 1 immunogenic peptide A01.
 KW MAGE 1; immunogenic peptide A01; cytotoxic C cells;
 KW in vitro activation; cancer; AIDS; bacterial infections; malaria;

KW fungal infections; tuberculosis; hepatitis.
 OS Homo sapiens.
 PN W09504817-A.
 PD 16-FEB-1995.
 PR 01-AUG-1994; U08672.
 PF 06-AUG-1993; US-103401.
 PA (CYTE-) CYTEL CORP.
 PI Cellis E, Kubo R, Serra H, Tsai V, Wentworth P;
 DR WPI: 95-090895/12.
 PT In vitro activation of cytotoxic T cells for selected killing of
 target cells - for treating e.g. cancer, AIDS, hepatitis etc. by
 incubating them with antigen presenting cells loaded with
 appropriate immunogenic peptide
 PS Example 3; Page 38; 53pp; English.
 CC R65109-R65145 are immunogenic peptides, they are used in a new
 method for the in vitro activation of cytotoxic T cells (CTC).
 CC This is achieved by incubating the CTCs with antigen presenting
 cells loaded with an appropriate immunogenic peptide (e.g. one
 of the above peptides). By selecting the peptides used the
 following diseases and infections can be treated; cancer, AIDS,
 hepatitis, other viral and bacterial infections, malaria and
 tuberculosis.
 CC Sequence 9 AA;
 SQ

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 |||||
 DB 1 EADPTGHSY 9

RESULT 7
 ID R70909 standard; Protein: 309 AA.
 AC R70909;
 DT 09-OCT-1995 (first entry)
 DE Human melanoma antigen MAGE-1.
 KW Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
 KW HLA-restricted cytotoxic T-lymphocyte activity.
 OS Homo sapiens.
 PN W09504542-A.
 PD 16-FEB-1995.
 PR 02-AUG-1994; U08721.
 PF 06-AUG-1993; US-103623.
 PA (CYTE-) CYTEL CORP.
 PI Fikes JD, Livingston BD, Sette AD, Sidney JC;
 DR WPI: 95-090681/12.
 DR N-PSDB; Q85435.
 PT Human melanoma antigen, MAGE-1, peptide(s) - useful for
 stimulating immune response against melanoma
 PS Example 1; Fig 1; 59pp; English.
 CC Q85435 encodes R70909 human melanoma antigen MAGE-1, it was used
 to produce the C-terminal MAGE-1 peptides described in R70915 to
 R70969. These peptides are useful for defining epitopes that
 engender a HLA-restricted cytotoxic lymphocyte activity against
 MAGE-1 antigens. Compsns. containing these peptides can be
 administered, as a vaccine to patients susceptible to MAGE
 associated tumours, e.g. melanomas.
 CC Sequence 309 AA;
 SQ

Query Match 100.0%; Score 52; DB 1; Length 309;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 |||||
 DB 161 EADPTGHSY 169

RESULT 8
 ID R65112 standard; peptide: 9 AA.
 AC R65112;
 DT 06-OCT-1995 (first entry)
 DE MAGE 1 immunogenic peptide 161-169
 KW MAGE 1; immunogenic peptide 161-169; cytotoxic C cells;
 KW in vitro activation; cancer; AIDS; bacterial infections; malaria;
 KW fungal infections; tuberculosis; hepatitis.
 OS Homo sapiens.
 PN W09504817-A.
 PD 16-FEB-1995.
 PF 01-AUG-1994; U08672.
 PR 06-AUG-1993; US-103401.
 PA (CYTE-) CYTEL CORP.
 PI Cellis E, Kubo R, Serra H, Tsai V, Wentworth P;
 DR WPI: 95-090895/12.
 PT In vitro activation of cytotoxic T cells for selected killing of
 target cells - for treating e.g. cancer, AIDS, hepatitis etc. by
 incubating them with antigen presenting cells loaded with
 appropriate immunogenic peptide
 PS Example 3; Page 35; 53pp; English.
 CC R65109-R65145 are immunogenic peptides, they are used in a new
 method for the in vitro activation of cytotoxic T cells (CTC).
 CC This is achieved by incubating the CTCs with antigen presenting
 cells loaded with an appropriate immunogenic peptide (e.g. one
 of the above peptides). By selecting the peptides used the
 following diseases and infections can be treated; cancer, AIDS,
 hepatitis, other viral and bacterial infections, malaria and
 tuberculosis.
 CC Sequence 9 AA;
 SQ

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 |||||
 DB 1 EADPTGHSY 9

RESULT 9
 ID R75954 standard; Peptide: 9 AA.
 AC R75954;
 DT 06-MAR-1996 (first entry)
 DE Melanoma antigen (MAGE-1) epitope.
 KW MAGE-3; melanoma antigen; vaccine; immune response; immunogenic peptide;
 KW cytotoxic T lymphocyte response; CTL; melanoma; breast cancer; antibody.
 OS Homo sapiens.
 PN W09519783-A1.
 PD 27-JUL-1995.
 PF 25-JAN-1995; U01000.
 PR 25-JAN-1994; US-186266.
 PA (CYTE-) CYTEL CORP.
 PI Cellis E, Grey HM, Kubo RT, Sette A;
 DR WPI: 95-269270/35.
 PT Immunogenic peptide(s) that induce immune response to cancer cells
 - that express a MAGE-3 protein peptide epitope used in vaccines or
 adoptive immunotherapy to induce cytotoxic T lymphocytes
 PS Example; Page 33; 44pp; English.
 CC R75954 is derived from MAGE-1 protein. It was used to show the
 specificity of CTL response to MAGE-3 peptides shown in R75942-53.
 CC R75942 is derived from the sequence of the melanoma antigen (MAGE-3)
 protein and can be used to elicit a primary cytotoxic T lymphocyte
 response against cells expressing MAGE-3. Synthetic peptides R75945-53
 can be used therapeutically to elicit CTL responses to melanoma, breast,
 colon, prostate, or other cells which express proteins with this epitope.
 CC The peptides have specific HLA-A1 binding capacity.
 CC Sequence 9 AA;
 SQ

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 |||||
 DB 1 EADPTGHSY 9

RESULT 10

R82988
 ID R82988 standard; Peptide; 9 AA.
 AC R82988;
 DT 26-FEB-1996 (first entry)
 DE P815 antigenic peptide.
 KW P815 antigen; P1A antigen; cancer; vaccine.
 OS Synthetic.
 PN W09523874-Al.
 PD 08-SEP-1995.
 PF 23-FEB-1995; U02203.
 PR 01-MAR-1994; US-204727.
 PR 10-MAR-1994; US-209172.
 PR 01-SEP-1994; US-299849.
 PR 30-NOV-1994; US-346774.
 PA (LUDW-) LUDWIG INST CANCER RES.
 PI Boon-Falleur T, Brasseur F, Chomez P, De Plaen E;
 PI De Smet C, Gaugler B, Lethe B, Marchand M, Patard J;
 PI Szikora J, Van Den Eynde B, Van Derbruggen P, Weynants P;
 DR WPI; 95-320586/41.
 PT Determin. of cancerous condition(s) - using a nucleic acid as a
 PT primer to determine expression of a MAGE tumour rejection antigen
 PT precursor
 PS Example 13; Page 22; 12pp; English.
 CC Using the sequence of the P815A antigen precursor gene P1A
 CC (T0117), an antigenic peptide (R82988) which was A+B+ (i.e.
 CC characteristic of cells which express both A and B antigens) was
 CC produced. The peptide lysed PO.HTR cells in the presence of
 CC cytolytic T lymphocyte cell lines, and may be useful as a vaccine
 CC component.
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 |||||
 DB 1 EADPTGHSY 9

RESULT 11

R78824
 ID R78824 standard; peptide; 9 AA.
 AC R78824;
 DT 26-MAR-1996 (first entry)
 DE MAGE-1 cytotoxic T lymphocyte epitope.
 KW MAGE-1; cytotoxic T; CTL; epitope; helper T; HTL; lymphocyte;
 KW cell; viruses; parasites; tumours; antigens; disease prevention;
 KW treatment.
 OS Homo sapiens.
 PN W09522317-Al.
 PD 24-AUG-1995.
 PF 16-FEB-1995; U02121.
 PR 16-FEB-1994; US-197484.
 PA (CITE-) CITE CORP.
 PI Cellis E, Chesnut RW, Grey H, Sette AD, Vitiello MA;
 DR WPI; 95-302545/39.
 PT Compsn. inducing cytotoxic T lymphocyte response to pref. viral,
 PT bacterial, parasitic or tumour antigens - useful in the treatment
 PT and prevention of diseases associated with the antigen e.g.
 PT hepatitis B
 PS Disclosure; Page 17; 109pp; English.

CC A compsn. which induces a cytotoxic T lymphocyte (CTL) response to
 CC an antigen (Ag) in a mammal comprises, a CTL Ag response inducing
 CC peptide (i.e. R78824-R78853) and a lipid conjugated helper T cell
 CC inducing peptide. The compsn. induces a CTL response to bacterial,
 CC viral or tumour Ags, and is therefore useful in the treatment and
 CC prevention of diseases associated with the Ag.
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 |||||
 DB 1 EADPTGHSY 9

RESULT 12

R80620
 ID R80620 standard; Protein; 12 AA.
 AC R80620;
 DT 28-FEB-1996 (first entry)
 DE Immunogenic peptide of tumour rejection antigen (MAGE-1).
 KW Tumour rejection; antigen; MAGE-1; monoclonal antibody; Mab;
 KW diagnosis; immunoassay; cancer; immunogen; antisera.
 OS Homo sapiens.
 PN W09520974-Al.
 PD 10-AUG-1995.
 PR 05-JAN-1995; U00095.
 PR 01-FEB-1994; US-190411.
 PA (LUDW-) LUDWIG INST CANCER RES.
 PA (SLOK) SLOAN KETTERING INST CANCER RES.
 PA (SLOK) MEMORIAL-SLOAN-KETTERING CANCER CENT.
 PI Boon-falleur T, Chen Y, Garin-chesa P, Old LJ, Rettig WJ;
 DR WPI; 95-283606/37.
 PT New monoclonal antibody binding specifically to MAGE-1 - useful for
 PT diagnosis and monitoring of cancer, also new hybridomas, recombinant
 PT MAGE-1 and immunogenic peptide(s)
 PS Claim 12; Page 20; 33pp; English.
 CC A monoclonal antibody directed against the tumour rejection antigen
 CC (MAGE-1) can be used to detect MAGE-1 in samples by standard
 CC immunoassay methods for diagnosis and monitoring of cancer etc. The
 CC monoclonal antibody is designated MA454 and is produced by the
 CC hybridoma deposited as ATCC HB11540. The monoclonal antibody is
 CC specific for MAGE-1, having no reactivity for MAGE-2 or MAGE-3.
 CC Peptide fragments of MAGE-1 (see R80618-20) may be useful as
 CC immunogens for production of the monoclonal antibody and antisera.
 SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 1; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00047;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
 |||||
 DB 4 EADPTGHSY 12

RESULT 13

R83932
 ID R83932 standard; peptide; 9 AA.
 AC R83932;
 DT 05-JUN-1996 (first entry)
 DE MHC class I restricted antigenic peptide #2.
 KW MHC class I; antigen; MAGE; melanoma; breast cancer; bladder cancer;
 KW Titermax; cytotoxic T-lymphocyte; tumour; pathogenic disease; bacteria;
 KW parasite; human; animal.
 OS Synthetic.
 PN W09528958-Al.
 PD 02-NOV-1995.

PF 21-APR-1995; U04975.
 PR 22-APR-1994; US-233496.
 PA (SLOK) SLOAN KETTERING INST CANCER RES.
 PI DYALL R, Nikolic-Zugic J;
 DR WPI: 95-382848/49.
 PT Cytotoxic T-cell induction by MHC class I-restricted peptide in
 PT adjuvant - useful for treating tumours and bacterial or parasitic
 PT pathogenic diseases
 PS Claim 11; Page 38; 50pp; English.
 CC The sequences given in R83931-49 are MHC class I restricted 8-12
 CC amino acid antigenic peptides. This peptide is derived from MAGE
 CC and is present in melanoma, breast and bladder cancer. These
 CC peptides may be administered to a subject in combination with a
 CC suitable adjuvant, pref. Titermax (RTM), to induce cytotoxic T-
 CC lymphocytes. This method may be used in the treatment of a tumour
 CC or a pathogenic disease, esp. diseases of bacterial or parasitic
 CC origin, in humans and animals, e.g. monkeys, dogs, cows, horses, etc.
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
 | | | | | | | | |
 Db 1 EADPTGHSY 9

RESULT 14

R90692
 ID R90692 standard; peptide; 9 AA.
 AC R90692;
 DT 31-JUL-1996 (first entry)
 DE Human leukocyte antigen (HLA-A1) presented peptide M23-E.
 KW Human leukocyte antigen; HLA-A1; MAGE-1 derived;
 KW blood mononuclear cell; BMC; CD8-beta+ cell; cytolytic T cell;
 KW CTL cell; treatment; tumour cell; diagnosis; assay;
 KW presented peptide.
 OS Synthetic.
 PN W09535500-A1.
 PD 28-DEC-1995; U07559.
 PF 14-JUN-1995; US-261541.
 PR 17-JUN-1994; US-261541.
 PA (LUDW-) LUDWIG INST CANCER RES.
 PI Boon-Falleur T, Coulie P, Van Der Bruggen P;
 DR WPI: 96-058510/06.
 PT Prodn. of specific cytolytic T cell sub-populations - by contacting
 PT blood mononuclear cells with specific peptide(s) and a population of
 PT CD8-beta(+) cells
 PS Claim 5; Page 19; 25pp; English.
 CC The present peptide is the human leukocyte antigen (HLA-A1), MAGE-1
 CC derived presented peptide, M23-E. By contacting a sample of blood
 CC mononuclear cells (BMC) with the peptide (which binds directly to
 CC HLA-A1 mols. on the surface of the BMC) and CD8-beta+ cells (which
 CC stimulate peptide/HLA-A1 complex specific CD8-beta+ cells), a
 CC peptide/HLA-A1 complex specific cytolytic T (CTL) cell
 CC subpopulation can be obtained. The CTL cells obtained can be
 CC administered to a patient to treat tumour cell related conditions,
 CC and can be used in diagnostic methods, e.g. in assays for the
 SQ peptide/HLA-A1 complex.
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
 | | | | | | | | |
 Db 1 EADPTGHSY 9

RESULT 15

R99343
 ID R99343 standard; Protein; 9 AA.
 AC R99343;
 DT 22-APR-1997 (first entry)
 DE MAGE-1 nonapeptide.
 KW HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
 KW tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
 KW antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
 KW therapy.
 OS Homo sapiens.
 PN W09626214-A1.
 PD 29-AUG-1996.
 PF 01-FEB-1995; U01489.
 PR 23-FEB-1995; US-393273.
 PA (LUDW-) LUDWIG INST CANCER RES.
 PI Boon-Falleur T, De Plaen E, Gaugler B, Lurquin C;
 PI Romero P, Traversari C, Van Den Eynde B, Van Der Bruggen P;
 DR WPI: 96-402317/40.
 DR N-PSDB: T35408.

PT New nona-peptide(s) that bind to HLA molecule(s) and induce lysis -
 PT by specific cytolytic T cells, for diagnosis and treatment of
 PT tumours and to expand T cells in vitro.
 PS Example 4; Fig 4; 41pp; English.
 CC R99343-R99350 represent MAGE nonapeptides, based on the tumour rejection
 CC antigen region of the full length MAGE sequences. These peptides were
 CC used to design the nonapeptides of the invention (see R99337-R99342),
 CC which bind to a HLA molecule on a cell, and provoke lysis by cytolytic T
 CC cells (CTLs) specific for a complex of the HLA molecule and nonapeptide.
 CC The nonapeptides can be used diagnostically to identify tumours
 CC expressing a particular HLA molecule, or to identify cancer cells. The
 CC peptides can also be used therapeutically, to induce a CTL response to
 CC tumours (where the peptides are optionally coupled to tumour-specific
 CC antibodies), or to induce a response by CTLs that are otherwise inactive.
 CC The peptide sequences may also be used to expand specific CTLs in vitro
 CC for later return to the patient, such as for treating melanoma. Tumour
 CC cells can be identified by using DNA encoding the nonapeptides as probes.
 CC Non-human cells transformed with the HLA-A1 gene and a DNA sequence
 CC encoding one of the peptides, can be used to generate CTLs, or to detect
 CC the presence of CTLs in human samples. The non-human transformed cells,
 CC when polytransformed, are universal effector cells, and can be used in
 CC vaccines, or for treating melanoma or breast cancer.
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
 | | | | | | | | |
 Db 1 EADPTGHSY 9

Search completed: December 10, 1999, 14:27:26
 Job time: 11234 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 1999, 07:27:34 ; Search time 10.2 Seconds
(without alignments)
35.352 Million cell updates/sec

Title: US-08-819-669D-26
Perfect score: 52
Sequence: 1 EADPTGHSY 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR_50.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	280	JC2358	tumor-associated a
2	44	84.6	369	I38659	MAGE-10 antigen - h
3	43	82.7	315	I38668	MAGE-9 antigen - h
4	43	82.7	315	I38667	MAGE-8 antigen - h
5	42	80.8	319	I38660	MAGE-11 antigen -
6	37	71.2	555	1 RGASWA	regulatory protein
7	36	69.2	3396	1 A42551	genome polyprotein
8	36	69.2	314	JC2361	tumor-associated a
9	36	69.2	314	JC2360	tumor-associated a
10	36	69.2	925	2 PH1299	MAGE 5 protein - h
11	35	67.3	925	2 A39216	plasma cell membra
12	35	67.3	1033	2 S02168	type I site-specif
13	35	67.3	98	2 F70769	hypothetical prote
14	34	65.4	497	2 S33938	penton protein (II
15	34	65.4	370	2 S49008	fork head protein
16	34	65.4	1670	2 S71551	DNA-directed DNA p
17	34	65.4	197	2 A70832	hypothetical prote
18	34	65.4	878	2 S44543	hypothetical prote
19	33	63.5	488	1 S55874	sulfite oxidase (E
20	33	63.5	488	1 A53107	sulfite oxidase (E
21	33	63.5	372	2 S32381	lignin peroxidase
22	33	63.5	775	2 S57920	DNA-directed DNA p
23	33	63.5	2123	2 S55089	probable acetyl-Co
24	33	63.5	295	2 C56180	adhesion protein -
25	33	63.5	288	2 A56279	carbon-monoxide de
26	33	63.5	747	2 B47093	cellulase (EC 3.2.
27	33	63.5	597	2 S37849	DNA intrastand cr
28	33	63.5	1040	2 A38306	alpha-mannosidase
29	33	63.5	301	2 C71194	hypothetical prote
30	32	61.5	133	1 F45345	trans-regulatory s
31	32	61.5	640	1 2VBPT5	tail protein pb5 -
32	32	61.5	283	2 E59626	methylentetrahedr
33	32	61.5	533	2 S71778	calcium-dependent
34	32	61.5	669	2 I38029	matrix metalloprot
35	32	61.5	349	2 A28658	nitrilase (EC 3.5.
36	32	61.5	506	2 S62629	phosphoglucosylase
37	32	61.5	381	2 A37276	leukocyte elastase
38	32	61.5	379	2 A42421	leukocyte elastase
39	32	61.5	700	2 S38928	translation elonga

40 32 61.5 396 2 G69808 multidrug resistan
41 32 61.5 2248 2 A35938 profilaggrin - hum
42 32 61.5 317 2 JC2359 tumor-associated a
43 32 61.5 689 2 S66006 conserved hypotbet
44 32 61.5 2529 2 B64635 toxin-like outer m
45 32 61.5 2411 2 A46299 tyrosine kinase su

ALIGNMENTS

RESULT 1

JC2358
tumor-associated antigen MAGE-1 - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 04-Sep-1998
C:Accession: JC2358
R: Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
Biochem. Biophys. Res. Commun. 202, 549-555, 1994
A:Title: Cloning and analysis of MAGE-1-related genes.
A:Reference number: JC2358; MUID:943111935
A:Accession: JC2358
A:Molecule type: mRNA
A:Residues: 1-280 <DIN>
A:Experimental source: melanoma cell line DMI50
C:Genetics:
A:Gene: MAGE
C:Superfamily: tumor associated protein MAGE
F:161-169/Region: HLA-A1 binding #status predicted

Query Match 100.0%; Score 52; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.0099;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
Db 161 EADPTGHSY 169

RESULT 2

I38659
MAGE-10 antigen - human
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 04-Sep-1998
C:Accession: I38659
R: De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.;
Oor, T.
Immunogenetics 40, 360-369, 1994
A:Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE
A:Reference number: I38659; MUID:95012457
A:Accession: I38659
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-369 <RES>
A:Cross-references: EMBL:U10685; NID:9533510; PID:9533511
C:Genetics:
A:Gene: GDB:MAGEA10; MAGE10
A:Cross-references: GDB:331126
A:Map position: Xq28-Xq28
A:Introns: #status absent
C:Superfamily: tumor associated protein MAGE

Query Match 84.6%; Score 44; DB 2; Length 369;
Best Local Similarity 77.8%; Pred. No. 0.45;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
Db 193 EVDPTGHSF 201

RESULT 3
I38668
MAGE-9 antigen - human
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 04-Sep-1998
C:Accession: I38668
R:De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; Br
oon, T.
Immunogenetics 40, 360-369, 1994
A:Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE fam
A:Reference number: I38659; MUID:95012457
A:Accession: I38668
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-315 <RES>
A:Cross-references: EMBL:U10694; NID:g533527; PID:g533528
C:Genetics:
A:Gene: GDB:MAGEA9; MAGE9
A:Cross-references: GDB:331125
A:Map position: Xp21.3-Xp21.3
A:Introns: #status absent
C:Superfamily: tumor associated protein MAGE

Query Match 82.7%; Score 43; DB 2; Length 315;
Best Local Similarity 77.8%; Pred. No. 0.6;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
| | | | |
DB 167 EVDPAHSHY 175

RESULT 4
I38667
MAGE-8 antigen - human
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 04-Sep-1998
C:Accession: I38667
R:De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; Br
oon, T.
Immunogenetics 40, 360-369, 1994
A:Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE fam
A:Reference number: I38659; MUID:95012457
A:Accession: I38667
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-234 <RES>
A:Cross-references: EMBL:U10693; NID:g533525; PID:g533526
C:Genetics:
A:Gene: GDB:MAGEA8; MAGE8
A:Cross-references: GDB:331123
A:Map position: Xq28-Xq28
A:Introns: #status absent
C:Superfamily: tumor associated protein MAGE

Query Match 82.7%; Score 43; DB 2; Length 234;
Best Local Similarity 77.8%; Pred. No. 0.44;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
| | | | |
DB 171 EVDPAHSHY 179

RESULT 5
I38660
MAGE-11 antigen - human
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 04-Sep-1998
C:Accession: I38660
R:De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; Br

oon, T.
Immunogenetics 40, 360-369, 1994
A:Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE
A:Reference number: I38659; MUID:95012457
A:Accession: I38660
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-319 <RES>
A:Cross-references: EMBL:U10686; NID:g533512; PID:g533513
C:Genetics:
A:Gene: GDB:MAGEA11; MAGE11
A:Cross-references: GDB:331128
A:Map position: Xq28-Xq28
A:Introns: #status absent
C:Superfamily: tumor associated protein MAGE

Query Match 80.8%; Score 42; DB 2; Length 319;
Best Local Similarity 77.8%; Pred. No. 0.94;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
| | | | |
DB 171 EVDPTSHSY 179

RESULT 6
RGASWA
regulatory protein weta - Emericella nidulans
C:Species: Emericella nidulans, Aspergillus nidulans
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Oct-1997
C:Accession: A39665
R:Marshall, M.A.; Timberlake, W.E.
Mol. Cell. Biol. 11, 55-62, 1991
A:Title: Aspergillus nidulans weta activates spore-specific gene expression.
A:Reference number: A39665; MUID:91094871
A:Accession: A39665
A:Molecule type: DNA
A:Residues: 1-555 <MAR>
A:Cross-references: GB:M60528; GB:M35758; NID:g168108; PID:g168109
C:Comment: The products of the genes bria, abaa, and weta are required for activation
C:Genetics:
A:Gene: weta
C:Superfamily: regulatory protein weta
C:Keywords: transcription regulation

Query Match 71.2%; Score 37; DB 1; Length 555;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EADPTGHS 8
| | | | |
DB 109 EADATGHS 116

RESULT 7
A42551
genome polyprotein - dengue virus type 1 (strain Singapore S275/90)
N:Contains: capsid protein; envelope protein; membrane protein; nonstructural protein
a: nonstructural protein NS4b; nonstructural protein NS5
C:Species: dengue virus type 1
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-May-1998
C:Accession: A42551
R:Fu, J.; Tan, B.H.; Yap, E.H.; Chan, Y.C.; Tan, Y.H.
Virology 188, 953-958, 1992
A:Title: Full-length cDNA sequence of dengue type 1 virus (Singapore strain S275/90).
A:Reference number: A42551; MUID:92263809
A:Accession: A42551
A:Molecule type: genomic RNA
A:Cross-references: GB:M87512
C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology

Query Match 87.5%; Score 37; DB 1; Length 555;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EADPTGHS 8
| | | | |
DB 109 EADATGHS 116

C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;
 F:1-114/Product: capsid protein #status predicted <CA>
 F:115-281/Product: membrane protein precursor #status predicted <MEP>
 F:115-204/Domain: nonterminal signal sequence #status predicted <SIG>
 F:205-281/Product: membrane protein #status predicted <MEM>
 F:267-279/Domain: transmembrane #status predicted <TM1>
 F:282-774/Product: envelope protein #status predicted <ENV>
 F:753-769/Domain: transmembrane #status predicted <TM2>
 F:775-1127/Product: nonstructural protein NS1 #status predicted <NS1>
 F:1128-1344/Product: nonstructural protein NS2a #status predicted <N2A>
 F:1345-1474/Product: nonstructural protein NS2b #status predicted <N2B>
 F:1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>
 F:1668-1938/Domain: DEAD/H box helicase homology <DEAD>
 F:1668-1675/Region: nucleotide-binding motif A (P-loop)
 F:1755-1760/Region: nucleotide-binding motif B
 F:1759-1762/Region: DEAD motif
 F:2094-2243/Product: nonstructural protein NS4a #status predicted <N4A>
 F:2244-2492/Product: nonstructural protein NS4b #status predicted <N4B>
 F:2493-3396/Product: nonstructural protein NS5 #status predicted <NS5>
 F:183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.2%; Score 36; DB 1; Length 3396;
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EADPTGHS 8

|||||

Db 3383 ESDPKGHS 3390

RESULT 8

JC2361
 tumor-associated antigen MAGE-3 - human
 N:Alternate names: MAGE 3 protein
 C:Species: Homo sapiens (man)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 04-Sep-1998
 C:Accession: JC2361; PH1296; I38438
 R:Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
 Biochem. Biophys. Res. Commun. 202, 549-555, 1994
 A:Title: Cloning and analysis of MAGE-1-related genes.
 A:Reference number: JC2358; MUID:93018875
 A:Accession: JC2361
 A:Molecule type: mRNA
 A:Residues: 1-314 <DIN>
 A:Experimental source: melanoma cell line DM150
 R:Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van Pel.
 J. Exp. Med. 176, 1453-1457, 1992
 A:Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic
 A:Reference number: PH1294; MUID:93018875
 A:Accession: PH1296
 A:Molecule type: DNA
 A:Residues: 168-176 <TRA>
 R:Gaigler, B.; van den Eynde, B.; van der Bruggen, P.; Romero, P.; Gaforio, J.J.; De Pla
 J. Exp. Med. 179, 921-930, 1994
 A:Title: Human gene MAGE-3 codes for an antigen recognized on a melanoma by autologous c
 A:Reference number: I38438; MUID:94157413
 A:Accession: I38438
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-314 <RES>
 A:Cross-references: EMBL:U03735; NID:9468825; PID:9468826
 C:Genetics:
 A:Gene: MAGE-3
 C:Superfamily: tumor associated protein MAGE
 F:168-176/Region: HLA-A1 binding #status predicted

Query Match 69.2%; Score 36; DB 2; Length 314;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EADPTGHS 9

Db 168 EVDPIGHLY 176

RESULT 9

JC2360
 tumor-associated antigen, MAGE 6 - human
 N:Alternate names: melanoma antigen 6; tumor-associated antigen, MAGE-3b
 C:Species: Homo sapiens (man)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 04-Sep-1998
 C:Accession: JC2360; PH1301; I38665; G01445
 R:Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
 Biochem. Biophys. Res. Commun. 202, 549-555, 1994
 A:Title: Cloning and analysis of MAGE-1-related genes.
 A:Reference number: JC2358; MUID:93018875
 A:Accession: JC2360
 A:Molecule type: mRNA
 A:Residues: 1-314 <DIN>
 A:Experimental source: melanoma cell line DM150
 R:Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van P
 J. Exp. Med. 176, 1453-1457, 1992
 A:Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic
 A:Reference number: PH1294; MUID:93018875
 A:Accession: PH1301
 A:Molecule type: DNA
 A:Residues: 168-176 <TRA>
 R:De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.;
 Immunogenetics 40, 360-369, 1994
 A:Title: Structure, chromosomal localization, and expression of 12 genes of the MAGE
 A:Reference number: I38659; MUID:95012457
 A:Accession: I38665
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-314 <RES>
 A:Cross-references: EMBL:U10691; NID:9533522; PID:9533523
 R:Fenton, R.G.
 submitted to the EMBL Data Library, June 1994
 A:Reference number: G07126
 A:Accession: G01445
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-314 <FEN>
 A:Cross-references: EMBL:U10339; NID:9499121; PID:9499122
 C:Genetics:
 A:Gene: GDB:MAGEA6; MAGE6
 A:Cross-references: GDB:331121
 Z:Map position: Xq28-Xq28
 A:Introns: #status absent
 C:Superfamily: tumor associated protein MAGE
 F:168-176/Region: HLA-A1 binding #status predicted

Query Match 69.2%; Score 36; DB 2; Length 314;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EADPTGHS 9

|||||

Db 168 EVDPIGHVY 176

RESULT 10

PH1299
 MAGE 5 protein - human (fragment)
 N:Alternate names: MAGE 51 protein
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 03-Aug-1998
 C:Accession: PH1299; PH1300
 R:Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van P
 J. Exp. Med. 176, 1453-1457, 1992
 A:Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic
 A:Reference number: PH1294; MUID:93018875

A.Accession: PH1299
A.Molecule type: DNA
A.Residues: 1-9 <TRA>
A.Accession: PH1300
A.Molecule type: DNA
A.Residues: 1-9 <TR2>

Query Match 69.2%; Score 36; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.2e+05;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
||||| :|
DB 1 EADPTNTY 9

RESULT 11
A39216

plasma cell membrane glycoprotein PC-1 - human
N:Contains: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.4.1)
C:Species: Homo sapiens (man)
C:Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 29-Aug-1997
C:Accession: A39216; S21706; S23587; S51030
R:Buckley, M.F.; Loveland, K.A.; McKinstry, W.J.; Garson, O.M.; Goding, J.W.
J. Biol. Chem. 265, 17506-17511, 1990
A:Title: Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human molecule, and
A:Reference number: A39216; MUID:91009202
A:Accession: A39216

A>Status: Preliminary
A.Molecule type: mRNA
A.Residues: 1-925 <BUC>
A:Cross-references: GB:J05654

R:Funakoshi, I.; Kato, H.; Horie, K.; Yano, T.; Hori, Y.; Kobayashi, H.; Inoue, T.; Suzuki, A.; Biochem. Biophys. 295, 180-187, 1992
A:Title: Molecular cloning of cDNAs for human fibroblast nucleotide pyrophosphatase.
A:Reference number: S21706; MUID:92246539
A:Accession: S21706

A>Status: not compared with conceptual translation

A.Molecule type: mRNA
A.Residues: 1-925 <FUNI>

A:Accession: S23587

A.Molecule type: protein
A.Residues: 116-121; 247-271, 'X', 273-275; 279-280, 'X', 282-283; 303-316; 362-364; 449-465; 482-

A>Note: It is uncertain whether Met-1 or Met-53 is the initiator
R:Belli, S.I.; Goding, J.W.

Eur. J. Biochem. 226, 433-443, 1994
A:Title: Biochemical characterization of human PC-1, an enzyme possessing alkaline phosphatase activity

A:Reference number: S51030
A:Accession: S51030

A>Status: Preliminary

A.Molecule type: mRNA

A.Residues: 1-80 <BEL>

C:Genetics:

A:Gene: GDB:PDNFI; M6S1; NPFS

A:Cross-references: GDB:132615; OMIM:173335

A:Map position: 6q22-6q23

C:Superfamily: nucleotide pyrophosphatase; somatomedin B homology

C:Keywords: glycoprotein; phosphoric diester hydrolase; transmembrane protein

F:104-144/Domain: somatomedin B homology <SBH1>

F:145-188/Domain: somatomedin B homology <SBH2>

F:179-285, 341, 477-578, 585, 643, 700, 731, 748/Binding site: carbohydrate (Asn) (covalent) #S

F:254/Active site: Thr (covalent substrate-binding) #status predicted

Query Match 67.3%; Score 35; DB 2; Length 925;
Best Local Similarity 66.7%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
| | | :| | |

Db 374 EPDSSGHSY 382

RESULT 12

S02168

type I site-specific deoxyribonuclease (EC 3.1.21.3) EcoRI24/3 chain hsdR - Escherichia coli
N:Alternate names: type I restriction enzyme EcoRI24/3 chain hsdR

C:Species: Escherichia coli

C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Sep-1997

C:Accession: S02168

R:Price, C.; Lingner, J.; Bickle, T.A.

J. Mol. Biol. 205, 115-125, 1989

A:Title: Basis for changes in DNA recognition by the EcoRI24 and EcoRI24/3 type I DNA

A:Reference number: S02166; MUID:89178628

A:Accession: S02168

A.Molecule type: DNA

A.Residues: 1-1033 <PRI>

A:Cross-references: EMBL:X13145; NID:g388978; PID:g41750

C:Genetics:

A:Gene: hsdR

A:Genome: plasmid

C:Keywords: DNA binding; hydrolase

Query Match 67.3%; Score 35; DB 2; Length 1033;
Best Local Similarity 66.7%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
| | | | | | |

Db 22 KAEPTGDSY 30

RESULT 13

F70769

hypothetical protein Rv1322 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Jul-1998

C:Accession: F70769

R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: F70769

A>Status: Preliminary; nucleic acid sequence not shown; translation not shown

A.Molecule type: DNA

A.Residues: 1-98 <COL>

A:Cross-references: GB:Z73902; GB:AL123456; NID:g3261576; PID:e245016; PID:g1340088

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv1322

Query Match 67.3%; Score 35; DB 2; Length 98;
Best Local Similarity 66.7%; Pred. No. 6.1;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
| | | | | | |

Db 24 EAGPDGHEY 32

RESULT 14

S33938

penton protein (III) - human adenovirus 12

C:Species: Mastadenovirus h12 (human adenovirus 12)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Sep-1997

C:Accession: S33938

R:Spiegel, J.

submitted to the EMBL Data Library, June 1993

A:Reference number: S33928
A:Accession: S33938
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-497 <SPR>
A:Cross-references: EMBL:X73487; NID:g313361; PID:g313372
C:Superfamily: adenovirus penton protein

Query Match 65.4%; Score 34; DB 2; Length 497;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
| | | | |
Db 310 ETPKGRSY 318

RESULT 15
S49008
fork head protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 17-Mar-1999
C:Accession: S49008; B56556
R:Lef, J.; Clement, J.H.; Oschwald, R.; Koester, M.; Knoechel, W.
Mech. Dev. 45, 117-126, 1994
A:Title: Spatial and temporal transcription patterns of the forkhead related XFD-2/XFD-2
A:Reference number: S49008; MUID:94257528
A:Accession: S49008
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-370 <LEF>
A:Cross-references: EMBL:X74315; NID:g511159; PID:g511160
R:Knoechel, S.; Lef, J.; Clement, J.; Klocke, B.; Hille, S.; Koster, M.; Knoechel, W.
Mech. Dev. 38, 157-165, 1992
A:Title: Activin A induced expression of a fork head related gene in posterior chordames
A:Reference number: A56556; MUID:93041288
A:Accession: B56556
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 118-228 <KNO>
A:Experimental source: gastrula
A:Note: sequence extracted from NCBI backbone (NCBIP:118173)
C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
F:127-218/Domain: fork head DNA-binding domain homology <FHD>

Query Match 65.4%; Score 34; DB 2; Length 370;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EADPTGHS 8
| | | | |
Db 257 ELSPTGHS 264

Search completed: December 11, 1999, 07:55:08
Job time: 1654 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 1999, 19:03:56 ; Search time 9.88 Seconds
(without alignments)
25.750 Million cell updates/sec

Title: US-08-819-669D-26

Perfect score: 52

Sequence: 1 EADPTGHSY 9

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt_37.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	309	1 MAG1_HUMAN	P43355 homo sapien
2	44	84.6	369	1 MAG8_HUMAN	P43363 homo sapien
3	43	82.7	234	1 MAG8_HUMAN	P43361 homo sapien
4	43	82.7	315	1 MAG9_HUMAN	P43362 homo sapien
5	42	80.8	319	1 MAGB_HUMAN	P43364 homo sapien
6	37	71.2	555	1 WETA_EMENI	P22022 emericella
7	36	69.2	314	1 MAG3_HUMAN	P43357 homo sapien
8	36	69.2	314	1 MAG6_HUMAN	P43360 homo sapien
9	36	69.2	3396	1 POLG_DEN1S	P33478 d genome po
10	35	67.3	346	1 MGB4_HUMAN	O15481 homo sapien
11	35	67.3	873	1 PC1_HUMAN	P22413 homo sapien
12	35	67.3	1033	1 TIR1_ECOLI	P10486 escherichia
13	35	67.3	98	1 Y037_MYCTU	Q10635 mycobacteri
14	34	65.4	497	1 PEN3_ADE12	P36716 human adeno
15	34	65.4	878	1 YB9X_YEAST	P38149 saccharomyc
16	33	63.5	1523	1 DPOL_THFEM	P74918 thermococcu
17	33	63.5	775	1 DPOL_THFEM	Q56366 thermococcu
18	33	63.5	747	1 GUND_CELFI	P50400 cellulomona
19	33	63.5	2273	1 HPA1_YEAST	P32874 saccharomyc
20	33	63.5	597	1 IXR1_YEAST	P33417 saccharomyc
21	33	63.5	372	1 LIGC_TRAVE	P20013 trametres ve
22	33	63.5	1040	1 MAN1_RAT	P21139 rattus norv
23	33	63.5	488	1 SUOX_HUMAN	P51687 homo sapien
24	33	63.5	488	1 SUOX_RAT	Q07116 rattus norv
25	33	63.5	503	1 VP57_BDV	P32638 borna disea
26	32	61.5	399	1 CAQS_MOUSE	O09165 mus musculu
27	32	61.5	395	1 CAQS_RABIT	P07221 oryctolagus
28	32	61.5	416	1 CC3_CANAL	P39826 candida alb
29	32	61.5	669	1 COGU_HUMAN	P51511 homo sapien
30	32	61.5	2411	1 DAB_DROME	P38081 drosophila
31	32	61.5	1002	1 DOR_DROME	Q43314 drosophila
32	32	61.5	700	1 EFG_AQUY	P46211 aquifex pyr
33	32	61.5	283	1 FOLD_BACSU	P54382 bacillus su
34	32	61.5	1360	1 GLI1_XENLA	Q91690 xenopus lae
35	32	61.5	1597	1 GTF1_STRDO	P11001 streptococc
36	32	61.5	1592	1 GTF2_STRDO	P27470 streptococc
37	32	61.5	379	1 ILEU_HORSE	P05619 equus cabal
38	32	61.5	280	1 LE76_BRANA	P13934 brassica na
39	32	61.5	317	1 MAG4_HUMAN	P43358 homo sapien
40	32	61.5	349	1 NRLB_KLEPN	P10045 klebsiella
41	32	61.5	269	1 RAD_HUMAN	P55042 homo sapien
42	32	61.5	268	1 RAD_RAT	P35043 rattus norv
43	32	61.5	133	1 REV_CAEOV	P33460 caprine art

ALIGNMENTS

RESULT 1

MAG1_HUMAN

ID MAG1_HUMAN STANDARD; PRT; 309 AA.
AC P43355; 000346;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 1 (MAGE-1 ANTIGEN) (ANTIGEN M22-E).
GN MAGE1 OR MAGE1 OR MAGE1A.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92086861.
RA VAN DER BRUGEN P., TRAVERSARI C., CHOMEZ P., LURQUIN C., DE PLAEN E.,
VAN DEN EYNDE B., KNUTH A., BOON T.;
RT "A gene encoding an antigen recognized by cytolytic T lymphocytes on
a human melanoma.";
RL SCIENCE 254:1643-1647(1991).
[2]
RP SEQUENCE FROM N.A.
RX TISSUE-SKIN;
RA MEDLINE; 94311935.
RT DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RL "Cloning and analysis of MAGE-1-related genes.";
RN BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
[3]
RP SEQUENCE FROM N.A.
RA GLOECKNER G., RUMP A., NORDSTIEK G., HINZMANN B., KIOSCHIS P.,
HEISS N., POUSTKA A., BAUER D., DRESCHER B., KNOB A., ROSENTHAL A.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
[4]
RP MUTAGENESIS.
RX TISSUE-BLOOD;
RA MEDLINE; 94157413.
RA GAUGLER B., VAN DEN EYNDE B., VAN DER BRUGEN P., ROMERO P.,
GAFORIO J.J., DE PLAEN E., LETHE B., BRASSEUR F., BOON T.;
RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
autologous cytolytic T lymphocytes.";
RL J. EXP. MED. 173:921-930(1994).
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CYTOLYTIC T LYMPHOCYTES.
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG,
CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
FOR TESTES. NEVER EXPRESSED IN KIDNEY TUMORS, LEUKEMIAS AND
LYMPHOMAS.
CC -!- POLYMORPHISM: THE VARIANT AT POSITION 32 LIKELY REPRESENTS A
POLYMORPHISM OF THE MAGE-1 GENE.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; M77481; G416115; -;
CC EMBL; U82672; G2078527; -;
CC MIN; 300016; -;
CC ANTIGEN; MULTIGENE FAMILY; POLYMORPHISM; TUMOR ANTIGEN.

```

FT VARIANT      32      32      T -> A.
FT DOMAIN      33      36      POLY-SER.
FT MUTAGEN     163     163     D->A: ABOLISHES HLA-A1 BINDING.
FT MUTAGEN     169     169     Y->A: ABOLISHES HLA-A1 BINDING.
FT CONFLICT    72      72      R -> Q (IN REF. 3).
SQ SEQUENCE    309 AA: 34342 MW; E6CB1300 CRC32;

Query Match      100.0%; Score 52; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.0091;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
Db 161 EADPTGHSY 169

RESULT 2
MAGA_HUMAN
ID MAGA_HUMAN STANDARD; PRT; 369 AA.
AC P43363;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 10 (MAGE-10 ANTIGEN).
GN MAGEA10 OR MAGE10.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA BRASSEUR R., CHOMEZ P., DE BACKER O., CAENEVEE W., BOON T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL IMMUNOGENETICS 40:360-369(1994).
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG,
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10685; G53351;
KW ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT DOMAIN 54 62 POLY-SER.
SQ SEQUENCE 369 AA: 40766 MW; D11E1870 CRC32;

Query Match      84.6%; Score 44; DB 1; Length 369;
Best Local Similarity 77.8%; Pred. No. 0.36;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
Db 193 EVDPTGHSF 201

RESULT 3
MAG8_HUMAN
ID MAG8_HUMAN STANDARD; PRT; 234 AA.
AC P43361;

Query Match      82.7%; Score 43; DB 1; Length 234;
Best Local Similarity 77.8%; Pred. No. 0.34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
Db 171 EVDPTGHSY 179

RESULT 4
MAG9_HUMAN
ID MAG9_HUMAN STANDARD; PRT; 315 AA.
AC P43362; Q92910;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 9 (MAGE-9 ANTIGEN).
GN MAGEA9 OR MAGE9.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA BRASSEUR R., CHOMEZ P., DE BACKER O., CAENEVEE W., BOON T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL IMMUNOGENETICS 40:360-369(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA TIMMS K.M., BONDESON M.L., ANSARI-LARI M.A., LAGERSTEDT K.,

```

RA NELSON D.L., PETERSSON U., GIBBS R.A.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U10694; G533528; -
DR EMBL; U66083; G1519285; -
KW ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT DOMAIN 34 37 POLY-GLU.
FT DOMAIN 87 90 POLY-GLU.
SQ SEQUENCE 315 AA: 35088 MW; 7DC3228E CRC32;

Query Match 82.7%; Score 43; DB 1; Length 315;
Best Local Similarity 77.8%; Pred. No. 0.47;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
| | | | |
Db 167 EVDPAHGSY 175

RESULT 5
MAGB_HUMAN
ID MAGB_HUMAN STANDARD; PRT; 319 AA.
AC P43364;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 11 (MAGE-11 ANTIGEN).
GN MAGE11 OR MAGE11.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95012457.
RA DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA BRASSEUR R., CHOMEZ P., DE BACKER O., CAVEENE W., BOON T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family".
RL IMMUNOGENETICS 40:360-369(1994).
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

CC
DR EMBL; U10686; G533513; -
KW ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
SQ SEQUENCE 319 AA; 35536 MW; E3DBEDEF CRC32;

Query Match 80.8%; Score 42; DB 1; Length 319;
Best Local Similarity 77.8%; Pred. No. 0.73;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
| | | | |
Db 171 EVDFTSHSY 179

RESULT 6
WETA_EMENI
ID WETA_EMENI STANDARD; PRT; 555 AA.
AC P22022;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE REGULATORY PROTEIN WETA.
GN WETA.
OS EMERICELLA NIDULANS (ASPERGILLUS NIDULANS).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PLECTOMYCETES;
OC EUROTIALES; TRICHOCOMACEAE; EMERICELLA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91094871.
RA MARSHALL M.A., TIMBERLAKE W.E.;
RT "Aspergillus nidulans weta activates spore-specific gene expression".
RL MOL. CELL. BIOL. 11:55-62(1991).
CC -!- FUNCTION: RESPONSIBLE FOR ACTIVATING A SET OF GENES WHOSE PRODUCTS
CC MAKE UP THE FINAL TWO CONIDIAL WALL LAYERS OR DIRECT THEIR
CC ASSEMBLY AND THOUGH THIS ACTIVITY IS RESPONSIBLE FOR ACQUISITION
CC OF SPORE DORMANCY.
CC -!- FUNCTION: BRLA, ABAA & WETA ARE PIVOTAL REGULATORS OF CONIDIOPHORE
CC DEVELOPMENT AND CONIDIUM MATURATION. THEY ACT INDIVIDUALLY AND
CC TOGETHER TO REGULATE THEIR OWN EXPRESSION AND THAT OF NUMEROUS
CC OTHER SPORULATION-SPECIFIC GENES.
CC -!- DEVELOPMENTAL STAGE: THE WETA GENE IS ACTIVATED ONLY DURING
CC CONIDIOPHORE DEVELOPMENT, AND ITS MRNA ACCUMULATES PREFERENTIALLY
CC IN MATURE CONIDIA.
CC -!- DOMAIN: HAS AN ACIDIC N-TERMINUS (AA 1-52) FOLLOWED BY A SER-,
CC THR-, PRO-RICH DOMAIN (AA 125-233) AND A BASIC C-TERMINUS (AA
CC 461-555).
CC -!- SIMILARITY: TO P.CHRYSOGENUM WETA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M60528; G168109; -
DR PIR; A39665; RGASWA.
KW DEVELOPMENTAL PROTEIN; CONIDIATION; TRANSCRIPTION REGULATION;
KW ACTIVATOR.
SQ SEQUENCE 555 AA; 60275 MW; 5B4F6B20 CRC32;

Query Match 71.2%; Score 37; DB 1; Length 555;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EADPTGHS 8
| | | | |
Db 109 EADATGHS 116

```

RESULT 7
MAG3_HUMAN
ID MAG3_HUMAN STANDARD; PRT: 314 AA.
AC P43357;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 3 (MAGE-3 ANTIGEN) (ANTIGEN M22-D).
GN MAGE3 OR MAGE3
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94157413.
RA GAUGLER B., VAN DEN ENDE B., VAN DER BRUGGEN P., ROMERO P.,
RA GAFORIO J.J., DE PLAEN E., LETHE B., BRASSEUR F., BOON T.;
RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
RT autologous cytolytic T lymphocytes."
RL J. EXP. MED. 179:921-930(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94311935.
RA DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RT "Cloning and analysis of MAGE-1-related genes."
RL BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
CC -!- FUNCTION: NOT KNOWN. THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC CYTOLYTIC T LYMPHOCYTES.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA. NEVER EXPRESSED IN KIDNEY TUMORS,
CC LEUKEMIAS AND LYMPHOMAS.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U03735; G468826;
DR DOMAIN 40 43 POLY-SER.
KW ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT MUTAGEN 170 170 D->A: ABOLISHES HLA-A1 BINDING.
FT MUTAGEN 176 176 Y->A: ABOLISHES HLA-A1 BINDING.
SQ SEQUENCE 314 AA; 34747 MW; AC557A64 CRC32;

Query Match 69.2%; Score 36; DB 1; Length 314;
Best Local Similarity 66.7%; Pred. No. 9.8;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
DB 168 EVDPIGHV 176

RESULT 8
MAG6_HUMAN
ID MAG6_HUMAN STANDARD; PRT: 314 AA.
AC P43360;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN 6 (MAGE-6 ANTIGEN) (MAGE3B).
GN MAGE6 OR MAGE6
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94157413.
RA GAUGLER B., VAN DEN ENDE B., VAN DER BRUGGEN P., ROMERO P.,
RA GAFORIO J.J., DE PLAEN E., LETHE B., BRASSEUR F., BOON T.;
RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
RT autologous cytolytic T lymphocytes."
RL J. EXP. MED. 179:921-930(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94311935.
RA DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RT "Cloning and analysis of MAGE-1-related genes."
RL BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
CC -!- FUNCTION: NOT KNOWN. THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC CYTOLYTIC T LYMPHOCYTES.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA. NEVER EXPRESSED IN KIDNEY TUMORS,
CC LEUKEMIAS AND LYMPHOMAS.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U03735; G468826;
DR DOMAIN 40 43 POLY-SER.
KW ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT MUTAGEN 170 170 D->A: ABOLISHES HLA-A1 BINDING.
FT MUTAGEN 176 176 Y->A: ABOLISHES HLA-A1 BINDING.
SQ SEQUENCE 314 AA; 34747 MW; AC557A64 CRC32;

Query Match 69.2%; Score 36; DB 1; Length 314;
Best Local Similarity 66.7%; Pred. No. 9.8;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
DB 168 EVDPIGHV 176

RESULT 9
POLG_DENIS
ID POLG_DENIS STANDARD; PRT: 3396 AA.
AC P33478;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE GENOME POLYPEPTIDE [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX
DE PROTEIN (ENVELOPE PROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL
DE PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3); RNA-DIRECTED
DE RNA POLYMERASE (EC 2.7.7.48) (NS5)].
OS DENGUE VIRUS TYPE 1 (STRAIN SINGAPORE S275/90).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;

```

```

OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95012457.
RA DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA DE SNET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LUQUIN C.,
RA BRASSEUR R., CHOMEZ P., DE BACKER O., CAVEENE W., BOON T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family."
RL IMMUNOGENETICS 40:360-369(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94311935.
RA DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RT "Cloning and analysis of MAGE-1-related genes."
RL BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95369706.
RA IMAI Y., SHICHIGO S., YAMADA A., KATAYAMA T., YANO H., ITOH K.;
RT "Sequence analysis of the MAGE gene family encoding human tumor-
RT rejection antigens."
RL GENE 160:287-290(1995).
CC -!- FUNCTION: NOT KNOWN. THOUGH MAY PLAY A ROLE IN TUMOR
CC OR ASPECTS OF TUMOR PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY TO
CC MAGE-3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10691; G533523;
DR EMBL; U10339; G499122;
DR EMBL; D32076; G1125016;
KW ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT DOMAIN 40 43 POLY-SER.
SQ SEQUENCE 314 AA; 34891 MW; B7125E97 CRC32;

Query Match 69.2%; Score 36; DB 1; Length 314;
Best Local Similarity 66.7%; Pred. No. 9.8;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
DB 168 EVDPIGHV 176

RESULT 9
POLG_DENIS
ID POLG_DENIS STANDARD; PRT: 3396 AA.
AC P33478;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE GENOME POLYPEPTIDE [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX
DE PROTEIN (ENVELOPE PROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL
DE PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3); RNA-DIRECTED
DE RNA POLYMERASE (EC 2.7.7.48) (NS5)].
OS DENGUE VIRUS TYPE 1 (STRAIN SINGAPORE S275/90).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;

```

OC FLAVIVIRUS.
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE; 92263809.
RA FU J., TAN B.H., YAP E.H., CHAN Y.C., TAN Y.H.;
RT "Full-length cDNA sequence of dengue type 1 virus (Singapore strain
RL S275/90).";
RL VIROLOGY 188:953-958(1992).
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF THREE PROTEINS:
CC PROTEIN PM, PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A
CC COMPLEX OF PROTEIN C AND MRNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M87512; -; NOT_ANNOTATED_CDS.
CC PIR; A42551; A42551.
CC PFAM; PF00869; Flavi_glycoprot; 1.
CC PFAM; PF00948; Flavi_NS1; 1.
CC PFAM; PF00949; Flavi_helicase; 1.
CC PFAM; PF00972; Flavi_NS5; 1.
CC PFAM; PF01002; Flavi_NS2B; 1.
CC PFAM; PF01003; Flavi_capsid; 1.
CC PFAM; PF01004; Flavi_M; 1.
CC PFAM; PF01005; Flavi_NS2A; 1.
CC HSP; P14336; 15VB.
CC POLYPROTEIN; GLYCOPROTEIN; TRANSFERASE; RNA-DIRECTED RNA POLYMERASE;
CC CORE PROTEIN; COAT PROTEIN; ENVELOPE PROTEIN; HELICASE; ATP-BINDING;
CC TRANSMEMBRANE; NONSTRUCTURAL PROTEIN.
CC TRANS CHAIN 1 114
CC PROPEP 115 205
CC CHAIN 206 280
CC CHAIN 281 774
CC CHAIN 775 1127
CC CHAIN 1128 1344
CC CHAIN 1345 1474
CC CHAIN 1475 2093
CC CHAIN 2094 2243
CC CHAIN 2244 2492
CC CHAIN 2493 3396
CC NP_BIND 1668 1675
CC SITE 1759 1762
CC TRANSMEM 267 279
CC TRANSMEM 753 769
CC DISULFID 283 310
CC DISULFID 340 396
CC DISULFID 354 385
CC DISULFID 372 401
CC DISULFID 465 565
CC DISULFID 582 613
CC CARBOHYD 183 183
CC CARBOHYD 347 347
CC CARBOHYD 433 433
CC SEQUENCE 3396 AA; 379558 MW; F538880D CRC32;

Query Match 69.2%; Score 36; DB 1; Length 3396;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 EADPTGHS 8
|:|:| |
Db 3383 ESDPKGHS 3390
|:|:| |
RESULT 10

MG84_HUMAN
ID MG84_HUMAN STANDARD; PRT; 346 AA.
AC O15481;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE MELANOMA-ASSOCIATED ANTIGEN B4 (MAGE-B4 ANTIGEN).
GN MAGE-B4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98110375.
RA LUDQUIN C., DE SMT C., BRASSEUR F., MUSCATELLI F., MARTELANGE V.,
RA DE PLAEN E., BRASSEUR R., MONACO A.P., BOON T.;
RT "Two members of the human MAGEB gene family located in Xp21.3 are
RT expressed in tumors of various histological origins";
RL GENOMICS 46:397-408(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA MUZY D., ARENSEN A.D., ADAMS C., BRUNDAGE E., BUNAC C., CARVELLI K.,
RA CHACKO J., CHEN J., DI W., DING Y., DUGAN S., DURBIN J., FORCUM J.,
RA GANESH R., GARCIA C., GOODMAN M., GORRELL J.H., HAYWOOD M.,
RA HERNANDEZ J., JACKSON L., JIN S., KAMPAL R., KARPATY S., KOVAR C.,
RA LEAL B., LI Y., LICHTARGE O., LIU W., LOGAN O., LU J., LY T.,
RA MARTINEZ C., OSWAL G., PEREZ L., RASHID N.D., ROWLAND K., SAVAGE L.,
RA SCHERER S.E., SHEN H., SIMON M., STOVALL K., TIMMS K.M., TODD J.,
RA VO O., WILLIAMSON A., WORLEY K.C., YU W., CHINAULT C., NELSON D.,
RA GIBBS R.A.;
RL SUBMITTED (OCT-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN TESTIS.
CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U93163; G2459682; -;
CC EMBL; AC005185; G3687199; -;
CC MIM; 300153; -;
CC KW ANTIGEN; MULTIGENE FAMILY.
CC SQ SEQUENCE 346 AA; 38923 MW; C6A5A407 CRC32;
Query Match 67.3%; Score 35; DB 1; Length 346;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 EADPTGHS 9
|:|:| |
Db 168 EVNPTTHSY 176
|:|:| |
RESULT 11
PC1_HUMAN
ID PC1_HUMAN STANDARD; PRT; 873 AA.
AC P22413;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1 (ALKALINE PHOSPHOTESTERASE I
DE (EC 3.1.4.1) / NUCLEOTIDE PYROPHOSPHATASE (EC 3.6.1.9) (NPase)).
GN PC1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE; 91009202.
RA BUCKLEY M.F., LOVELAND K.A., MCKINSTRY W.J., GARSON O.M., GODING J.W.;
RT "Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human
RT molecule, amino acid sequence, and chromosomal location.";
RL J. BIOL. CHEM. 265:17508-17511(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92246539.
RA FUNAKOSHI I., KATO H., HORIE K., YANO T., HORI Y., KOBAYASHI H.,
RA INOUE T., SUZUKI H., FUKUI S., TSUKAHARA M., KAJI T.,
RA YAMASHINA I.;
RT "Molecular cloning of cDNAs for human fibroblast nucleotide
RT pyrophosphatase";
RL ARCH. BIOCHEM. 295:180-187(1992).
CC -!- FUNCTION: MAY HAVE A ROLE IN THE REGULATION OF N-GLYCOSYLATION.
CC -!- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES
CC SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED
CC OLIGO-NUCLEOTIDES.
CC -!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN PLASMA CELLS AND ALSO IN A NUMBER
CC OF NON-LYMPHOID TISSUES, INCLUDING THE DISTAL CONVOLUTED TUBULE
CC OF THE KIDNEY, CHONDROCYTES, AND EPIDIDYMIS.
CC -!- SIMILARITY: CONTAINS 2 SOMATOMEDIN-B TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M57736; G189650; -
DR EMBL; D12485; G219945; -
DR EMBL; D12485; G219944; ALT_INIT.
DR PIR; A39216; A39216.
DR MIM; 173335; -
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
DR PFAM; PF01033; Somatomedin_B; 2.
KW GLYCOPROTEIN; TRANSMEMBRANE; DUPLICATION; SIGNAL-ANCHOR; HYDROLASE.
FT DOMAIN 1 24
FT TRANSMEM 25 45
FT DOMAIN 46 873
FT DOMAIN 52 92
FT DOMAIN 93 136
FT CARBOHYD 127 127
FT CARBOHYD 233 233
FT CARBOHYD 289 289
FT CARBOHYD 425 425
FT CARBOHYD 533 533
FT CARBOHYD 591 591
FT CARBOHYD 648 648
FT CARBOHYD 679 679
FT CARBOHYD 696 696
SQ SEQUENCE 873 AA; 99929 MW; 580583CD CRC32;

Query Match 67.3%; Score 35; DB 1; Length 873;
Best Local Similarity 66.7%; Pred. NO. 44;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
DB 322 EPDSSGHSY 330

RESULT 12
T1RL_ECOLI
ID T1RL_ECOLI STANDARD; PRT; 1033 AA.
AC P10486;
DT 01-JUL-1989 (REL. 11, CREATED)

DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE TYPE I RESTRICTION ENZYME ECOL124II R PROTEIN (EC 3.1.21.3).
GN HSR OR HSR.
OS ESCHERICHIA COLI.
OG PLASMID INCFIV R124/3.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89178628.
RA PRICE C., LINGNER J., BICKLE J., FIRMAN T.A., GLOVER S.W.;
RT "Basis for changes in DNA recognition by the Ecor124 and Ecor124/3
RT type I DNA restriction and modification enzymes.";
RL J. MOL. BIOL. 205:115-125(1989).
CC -!- FUNCTION: THE Ecor124/3 I ENZYME RECOGNIZES 5'GAA(N7)RTCG.
CC -!- FUNCTION: SUBUNIT R IS REQUIRED FOR BOTH NUCLEASE AND ATPASE
CC ACTIVITIES, BUT NOT FOR MODIFICATION.
CC -!- SUBUNIT: THE TYPE I RESTRICTION & MODIFICATION SYSTEM IS COMPOSED
CC OF THREE POLYPEPTIDES R, M AND S.
CC -!- TYPE I RESTRICTION AND MODIFICATION ENZYMES ARE COMPLEX. MULTI-
CC FUNCTIONAL SYSTEMS WHICH REQUIRE ATP, S-ADENOSYL METHIONINE AND
CC MG(2+) AS CO-FACTORS AND, IN ADDITION TO THEIR ENDONUCLEOLYTIC
CC AND METHYLASE ACTIVITIES, ARE POTENT DNA-DEPENDENT ATPASES.
CC -!- SIMILARITY: WITH ATPASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X13145; G41750; -
DR PIR; S02168; S02168.
DR REBASE; RB00748; Ecor124II.
KW PLASMID; RESTRICTION SYSTEM; HYDROLASE; DNA-BINDING; ATP-BINDING.
SQ SEQUENCE 1033 AA; 119656 MW; 9E98CC1 CRC32;

Query Match 67.3%; Score 35; DB 1; Length 1033;
Best Local Similarity 66.7%; Pred. NO. 52;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EADPTGHSY 9
DB 22 KAEPTGDSY 30

RESULT 13
Y037_MYCTU
ID Y037_MYCTU STANDARD; PRT; 98 AA.
AC Q10635;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE VERY HYPOTHETICAL 11.3 KD PROTEIN CY130.07.
GN MYC130.07.
OS MYCOBACTERIUM TUBERCULOSIS.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIACEAE;
OC ACTINOMYCETALES; CORYNEBACTERIACEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; 273902; E245016; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 98 AA; 11334 MW; 8E3C8C75 CRC32;

Query Match 67.3%; Score 35; DB 1; Length 98;
Best Local Similarity 66.7%; Pred. No. 4.5;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
| | | | |
Db 24 EAGPDGHEY 32

RESULT 14
PEN3_ADE12 STANDARD; PRT: 497 AA.
ID PEN3_ADE12
AC P36716;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PENTON PROTEIN (VIRION COMPONENT III) (PENTON BASE PROTEIN).
GN PIII.
OS HUMAN ADENOVIRUS TYPE 12.
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; ADENOVIRIDAE; MASTADENOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94076430.
RA SPRENGEL J., SCHMITZ B., HEUSS-NEITZEL D., ZOCK C., DOERFLER W.;
RT "Nucleotide sequence of human adenovirus type 12 DNA: comparative
functional analysis."
RL J. VIROL. 68:379-389(1994).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; X73487; G313372; -
DR PIR; S33938; S33938.
KW LATE PROTEIN.
SQ SEQUENCE 497 AA; 56393 MW; A5BEC571 CRC32;

Query Match 65.4%; Score 34; DB 1; Length 497;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9
| | | | |
Db 310 ETDPKGRSY 318

RESULT 15
YB9X_YEAST STANDARD; PRT: 878 AA.
ID YB9X_YEAST
AC P38149;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 98.1 KD TRP-ASP REPEATS CONTAINING PROTEIN IN PAF1-MRPL27
DE INTERGENIC REGION.
GN YBR281C OR YBR2018.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCETES.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE; 94378722.
RA HOLMSTROEM K., BRANDT T., KALLESOE T.;
RT "The sequence of a 32,420 bp segment located on the right arm of
RT chromosome II from Saccharomyces cerevisiae."
RL YEAST 10:47-62(1994).
CC -----
CC !- SIMILARITY: CONTAINS ? WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; X76053; G429126; -
DR EMBL; Z36150; G536726; -
DR PIR; S44543; S44543.
DR PIR; S39137; S39137.
DR PROSITE; PS00678; WD_REPEATS; 2.
KW HYPOTHETICAL PROTEIN; REPEAT; WD REPEAT.
SQ SEQUENCE 878 AA; 98070 MW; 093B13F4 CRC32;

Query Match 65.4%; Score 34; DB 1; Length 878;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 DPTGHSY 9
| | | | |
Db 796 DPTNHAY 802

Search completed: December 8, 1999, 23:56:43
Job time: 17567 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 1999, 00:13:00 ; Search time 15.25 Seconds
(without alignments)
36.321 Million cell updates/sec

Title: US-08-819-669d-26

Perfect score: 52

Sequence: 1 EADPTGHSY 9

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL10:

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	71.2	1032	11	Q61989 mus musculus
2	36	69.2	2457	12	Q41965 murine herp
3	35	67.3	1187	2	Q59278 cellulomona
4	35	67.3	131	2	O85701 streptomyc
5	35	67.3	330	11	Q60763 mus musculu
6	35	67.3	330	11	Q60761 mus musculu
7	35	67.3	320	11	O89006 mus musculu
8	35	67.3	325	11	O89010 mus musculu
9	34	65.4	1671	1	P77933 pyrococcus
10	34	65.4	197	2	O53701 mycobacteri
11	34	65.4	224	2	O69955 streptomyc
12	34	65.4	215	2	O87835 streptomyc
13	34	65.4	268	4	Q14969 homo sapien
14	34	65.4	1184	4	O75339 homo sapien
15	34	65.4	347	4	O75862 homo sapien
16	34	65.4	347	4	O00601 homo sapien
17	34	65.4	156	5	Q23961 drosophila
18	34	65.4	599	6	O19112 sus scrofa
19	34	65.4	272	11	O35830 rhizomys pr
20	34	65.4	267	11	O70217 mus musculu
21	34	65.4	342	11	O88737 mus musculu
22	34	65.4	320	11	O89009 mus musculu
23	34	65.4	906	12	Q9YU1 hemorhagic
24	34	65.4	370	13	Q91904 xenopus lae
25	34	63.5	295	1	O26704 methanobact
26	33	63.5	301	1	O59491 pyrococcus
27	33	63.5	1086	2	O69230 bacillus sp
28	33	63.5	307	2	O24828 acinetobact
29	33	63.5	288	2	Q51323 pseudomonas

ALIGNMENTS

RESULT 1

Q61989 ID Q61989 PRELIMINARY; PRT: 1032 AA.
AC Q61989; Q61740;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE INTEGRIN ALPHA-4 SUBUNIT PRECURSOR (INTEGRIN ALPHA-IV) (VLA-4) (CD49D)
DE (LYMPHOCYTE-Peyer's PATCH ADHESION MOLECULES ALPHA SUBUNIT) (LPAM)
DE ALPHA SUBUNIT)
GN ITGA4 OR VLA-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RF SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=0BA, AND NIH/SWISS; TISSUE=LEUKEMIA, AND FIBROBLAST;
RX MEDLINE: 95290094.
RA DE MEIRSMAN C., SCHOLLEN E., JASPERS M., ONGENA K., MATTHIJS G.,
RA MARYNEN P., CASSIMAN J.J.;
RT "Cloning and characterization of the promoter region of the murine
alpha-4 integrin subunit.";
RL DNA CELL BIOL. 13:743-754(1994).
RN [2]
RF SEQUENCE FROM N.A.
RC STRAIN=B6/CBA, AND NIH/SWISS; TISSUE=SPLEEN, AND FIBROBLAST;
RX MEDLINE: 96326295.
RA DE MEIRSMAN C., JASPERS M., SCHOLLEN E., CASSIMAN J.J.;
RL "The genomic structure of the murine alpha 4 integrin gene.";
RN [3]
RF DNA CELL BIOL. 15:595-603(1996).
RN [4]
RF SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=LIVER;
RA HASEGAWA M., FOOTE S.;
RL SUBMITTED (NOV-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]
RF SEQUENCE OF 134-307 FROM N.A.
RN [6]
RF TISSUE=MELANOMA;
RA ROUT U.K., ARMANI D.R.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [7]
RF FUNCTION: THE ALPHA-4 BETA-1 INTEGRIN MEDIATES ADHESION TO ONE OR
MORE DOMAINS WITHIN THE HEP1-11CS REGION OF FIBRONECTIN
INCLUDING THE ALTERNATIVELY SPLICED CS-1 REGION. VLA-4 ALSO
RECOGNIZES THE CELLULAR LIGAND VCAM-1 ON ACTIVATED ENDOTHELIAL
CELLS. IT TRIGGERS HOMOTYPIC AGGREGATION FOR MOST VLA-4-POSITIVE
LEUKOCYTE CELL LINES. VLA-4 MAY ALSO PARTICIPATE IN CYTOLYTIC
T-CELL INTERACTIONS WITH TARGET CELLS. IMPORTANT FOR CELL-CELL
ADHESION FUNCTIONS (BY SIMILARITY).
RN [8]
RF FUNCTION: THE ALPHA-4 BETA-7 HETERODIMER INTERACTS WITH THE
LYMPHOCYTE-Peyer's PATCH ADHESION MOLECULES VLA-4 AND IS
IMPORTANT FOR CELL-CELL ADHESION FUNCTIONS. VLA-4 MAY PLAY A
SUPPORTIVE ROLE IN THE HOMING OF LYMPHOCYTES TO THE PEYER'S

30 33 63.5 2273 3 042823 042823 saccharomyc
31 33 63.5 827 3 043048 043048 schizosacch
32 33 63.5 341 5 025028 025028 haemochth
33 33 63.5 341 10 043449 043449 glycine max
34 33 63.5 245 10 023713 023713 arabidopsis
35 33 63.5 246 10 081146 081146 arabidopsis
36 33 63.5 246 10 081147 081147 arabidopsis
37 33 63.5 106 10 092RPO 092RPO caragana ar
38 33 63.5 503 12 089857 089857 borna disea
39 33 63.5 503 12 065459 065459 borna disea
40 33 63.5 503 12 010394 010394 borna disea
41 33 63.5 503 12 010397 010397 borna disea
42 33 63.5 503 12 010399 010399 borna disea
43 33 63.5 503 12 010400 010400 borna disea
44 33 63.5 503 12 010403 010403 borna disea
45 33 63.5 353 12 088626 088626 borna disea

CC PATCHES.
CC -!- SUBUNIT: DIMER OF AN ALPHA AND BETA CHAIN. THE BETA CHAIN THAT
CC ASSOCIATES WITH ALPHA-4 IS EITHER BETA-1 OR BETA-7. THE ALPHA-4
CC BETA-1 HETERODIMER IS REFERRED TO AS VLA-4 OR LPAM-2. THE ALPHA-4
CC BETA-7 HETERODIMER IS REFERRED TO AS LPAM-1 OR THE PEYER'S PATCH
CC HOMING RECEPTOR. THE ALPHA CHAIN CAN SOMETIMES BE CLEAVED INTO TWO
CC NON-COVALENTLY ASSOCIATED FRAGMENTS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: THE VLA-4 INTEGRIN IS MAINLY EXPRESSED BY
CC LYMPHOID AND MYELOID CELLS. THE ALPHA-4 BETA-7 INTEGRIN IS
CC EXPRESSED BY PEYER'S PATCH HOMING CELLS.
CC -!- MISCELLANEOUS: THIS PROTEIN HAS 3 POTENTIAL CALCIUM-BINDING SITES.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC
DR EMBL; L20788; AAA97501.1; -
DR EMBL; U34800; AAB09630.1; -
DR EMBL; U34827; AAB09630.1; JOINED.
DR EMBL; U34828; AAB09630.1; JOINED.
DR EMBL; U34829; AAB09630.1; JOINED.
DR EMBL; U34830; AAB09630.1; JOINED.
DR EMBL; U34831; AAB09630.1; JOINED.
DR EMBL; U34832; AAB09630.1; JOINED.
DR EMBL; U34833; AAB09630.1; JOINED.
DR EMBL; U34834; AAB09630.1; JOINED.
DR EMBL; U34835; AAB09630.1; JOINED.
DR EMBL; U34763; AAB09630.1; JOINED.
DR EMBL; U34764; AAB09630.1; JOINED.
DR EMBL; U34765; AAB09630.1; JOINED.
DR EMBL; U34766; AAB09630.1; JOINED.
DR EMBL; U34767; AAB09630.1; JOINED.
DR EMBL; U34768; AAB09630.1; JOINED.
DR EMBL; U34769; AAB09630.1; JOINED.
DR EMBL; U34770; AAB09630.1; JOINED.
DR EMBL; U34797; AAB09630.1; JOINED.
DR EMBL; U34798; AAB09630.1; JOINED.
DR EMBL; U34799; AAB09630.1; JOINED.
DR EMBL; AF109136; AAC95388.1; -
DR EMBL; U97151; AAC05709.1; -
DR MGD; MGI:96603; ITGA4.
DR PFAM; PF00357; Integrin.A; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
KW Integrin; Cell adhesion; Glycoprotein; Transmembrane; Signal;
KW Extracellular matrix; Cytoskeleton; Calcium-binding; Polymorphism.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 1032 INTEGRIN ALPHA-4 SUBUNIT.
SQ SEQUENCE 1032 AA; 115085 MW; 27B379B1 CRC32;

Query Match 71.2%; Score 37; DB 11; Length 1032;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 PTGHSY 9
Db 29 PTGHSY 34

RESULT 2
O41965 PRELIMINARY; PRT; 2457 AA.
ID O41965;
AC O41965;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE TEGUMENT PROTEIN.
GN GAMMAHV. ORF64.
OS murine herpesvirus 68.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WUMS;
RX MEDLINE; 97366649.

RA VIRGIN H.W. IV, LATREILLE P., WAMSLEY P., HALLSWORTH K., WECK K.E.,
RA DAL CANTO A.J., SPECK S.H.;
RT "Complete sequence and genomic analysis of murine gammaherpesvirus
68";
RL J. Virol. 71:5894-5904(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-WUMS;
RA LATREILLE P., WAMSLEY P., WATERSTON R.H.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U97553; AAB66417.1; -
SQ SEQUENCE 2457 AA; 273535 MW; 8DC81BBD CRC32;

Query Match 69.2%; Score 36; DB 12; Length 2457;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DPTGHSY 9
Db 1531 DPTGHSY 1537

RESULT 3
O59278 PRELIMINARY; PRT; 1187 AA.
ID O59278;
AC O59278;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
DE ENDOXYLANASE (EC 3.2.1.8) (ENDO-1,4-BETA-XYLANASE)
DE (1,4-BETA-D-XYLAN XYLANOXYDROLASE).
GN XYN.
OS Cellulomonas.fim.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococineae; Cellulomonadaceae; Cellulomonas.
RN [1]
RP SEQUENCE OF 1-352 FROM N.A.
RX MEDLINE; 96245431.
RA CLARKE J.H., DAVIDSON K., GILBERT H.J., FONTES C.M., HAZLEWOOD G.P.;
RT "A modular xylanase from mesophilic Cellulomonas fimi contains the
RT same cellulose-binding and thermostabilizing domains as xylanases
RT from thermophilic bacteria";
RL FEMS Microbiol. Lett. 139:27-35(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA CLARKE J.H.;

RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
CC LINKAGES IN XYLANS.
DR EMBL; Z50866; CAA90745.1; -
DR PFAM; PF00331; Glyco_hydro_10; 1.
KW Xylan degradation; Hydrolase; Glycosidase.
SQ SEQUENCE 1187 AA; 125378 MW; 92B3994A CRC32;

Query Match 67.3%; Score 35; DB 2; Length 1187;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 DPTGHSY 9
Db 1047 DPTGHSY 1053

RESULT 4
O85701 PRELIMINARY; PRT; 131 AA.
ID O85701;
AC O85701;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE HYPOTHETICAL 14.2 KD PROTEIN.

OS Streptomyces lividans.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1326, AJ100;
 RA ALTENBUCHNER J.;
 RT "Amplifiable element AUD4 from Streptomyces lividans 66";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF072709; AAC25770.1;
 KW Hypothetical protein.
 SQ SEQUENCE 131 AA; 14187 MW; 8321BCE1 CRC32;

Query Match 67.3%; Score 35; DB 2; Length 131;
 Best Local Similarity 62.5%; Pred. No. 12;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ADPTGHSY 9

Db 110 SDPAGHSF 117

RESULT 5
 Q60763 PRELIMINARY; PRT; 330 AA.
 AC Q60763;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE MELANOMA ANTIGEN, RELATED SEQUENCE 2 (SMAGE-3 PROTEIN).
 GN MAGE-RS3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DBA/2; TISSUE=KIDNEY;
 RX MEDLINE; 96070435.
 RA DE BACKER O., VERHEYDEN A.M., MARTIN B., GODELAINE D., DE PLAEN E.,
 RA BRASSEUR R., AVNER P., BOON T.;
 RT "Structure, chromosomal location, and expression pattern of three
 mouse genes homologous to the human MAGE genes";
 RL Genomics 28:74-83(1995).
 DR EMBL; U19033; AAA86098.1;
 DR MGD; MGI:105109; MAGE-RS3.
 DR PFAM; PF01454; MAGE; 1.
 SQ SEQUENCE 330 AA; 35985 MW; 83AD4246 CRC32;

Query Match 67.3%; Score 35; DB 11; Length 330;
 Best Local Similarity 66.7%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9

Db 161 EIDPSTHSY 169

RESULT 6
 Q60761 PRELIMINARY; PRT; 330 AA.
 AC Q60761; Q60762;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE MELANOMA ANTIGEN RELATED SEQUENCES 1 AND 2
 (SMAGE-1 PROTEIN / SMAGE-2 PROTEIN).
 GN (SMAGE-RS1 OR SMAGE1) AND (SMAGE-RS2 OR SMAGE2).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=DBA/2; TISSUE=KIDNEY;
 RX MEDLINE; 96070435.
 RA DE BACKER O., VERHEYDEN A.M., MARTIN B., GODELAINE D., DE PLAEN E.,
 RA BRASSEUR R., AVNER P., BOON T.;
 RT "Structure, chromosomal location, and expression pattern of three
 mouse genes homologous to the human MAGE genes";
 RL Genomics 28:74-83(1995).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN TUMOURS OF VARIOUS HISTOLOGICAL
 CC TYPES BUT NOT IN NORMAL TISSUES EXCEPT TESTIS.
 CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
 DR EMBL; U19031; AAA86096.1; ALT_INIT.
 DR EMBL; U19032; AAA86097.1;
 DR MGD; MGI:105117; MAGE-RS2.
 DR PFAM; PF01454; MAGE; 1.
 KW Antigen; Tumor antigen.
 SQ SEQUENCE 330 AA; 35936 MW; 36D760C5 CRC32;

Query Match 67.3%; Score 35; DB 11; Length 330;
 Best Local Similarity 66.7%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9

Db 161 EIDPSTHSY 169

RESULT 7
 Q89006 PRELIMINARY; PRT; 320 AA.
 ID Q89006;
 AC Q89006;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE MAGEAL PROTEIN.
 GN MAGEAL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA DE PLAEN E., DE BACKER O., ARNAUD D., BONJEAN B., CHOMEZ P.,
 RA MARTELANGE V., AVNER P., BALDACCIO P., BABINET C., HWANG SY.,
 RA KNOWLES B., BOON T.;
 RT "A new family of mouse genes homologous to the human MAGE genes";
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ005525; CAA06579.1;
 DR PFAM; PF01454; MAGE; 1.
 SQ SEQUENCE 320 AA; 36096 MW; 914E8580 CRC32;

Query Match 67.3%; Score 35; DB 11; Length 320;
 Best Local Similarity 66.7%; Pred. No. 30;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EADPTGHSY 9

Db 155 EVDPSVHSY 163

RESULT 8
 Q89010 PRELIMINARY; PRT; 325 AA.
 ID Q89010;
 AC Q89010;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE MAGEA6 PROTEIN.
 GN MAGEA6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA DE PLAEN E., DE BACKER O., ARNAUD D., BONJEAN B., CHOMEZ P.,
 RA MARTELANGE V., AVNER P., BALDACCI P., BABINET C., HWANG SY.,
 RA KNOWLES B., BOON T.;
 RT "A new family of mouse genes homologous to the human MAGE genes.";
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ005530; CAA06584.1;
 DR PFAM: PF01454; MAGE.1;
 SQ SEQUENCE 325 AA; 36640 MW; 3B3022A3 CRC32;
 Query Match 67.3%; Score 35; DB 11; Length 325;
 Best Local Similarity 66.7%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EADPTGHSY 9
 I | | | | |
 Db 155 EVDPVSHSY 163
 RESULT 9
 ID P77933 PRELIMINARY; PRT: 1671 AA.
 AC P77933;
 DT 01-FEB-1997 (TEMBLrel. 02, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE DNA-DEPENDENT DNA POLYMERASE (EC 2.7.7.7)
 DE (DNA-DIRECTED DNA POLYMERASE) (DNA NUCLEOTIDYLTRANSFERASE
 DE (DNA-DIRECTED))
 GN POL.
 OS Pyrococcus sp.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KOD1;
 RX MEDLINE; 98027387.
 RA TAKAGI M., NISHIOKA M., KAKIHARA H., KITABAYASHI M., INOUE H.,
 RA KAKAKAMI B., OKA M., IMAOKA T.;
 RT "Characterization of DNA polymerase from Pyrococcus sp. strain KOD1
 RT and its application to PCR.";
 RL Appl. Environ. Microbiol. 63:4504-4510(1997).
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
 CC DIPHOSPHATE + DNA(N).
 DR EMBL: D29671; BAA06142.1;
 DR PFAM: PF00136; DNA_pol_B.3;
 KW Transferase; Nucleotidyltransferase; Endonuclease.
 FT CHAIN 407 766
 FT CHAIN 852 1388
 FT CHAIN 852 1388
 SQ SEQUENCE 1671 AA; 193432 MW; 33C27AEE CRC32;
 Query Match 65.4%; Score 34; DB 1; Length 1671;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EADPTGHSY 9
 I | | | | |
 Db 1616 EFDPTKHY 1624
 RESULT 10
 ID O53701 PRELIMINARY; PRT: 197 AA.
 AC O53701;
 DT 01-JUN-1998 (TEMBLrel. 06, Created)
 DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE HYPOTHETICAL 21.4 KD PROTEIN.
 GN MT036.01C.

OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RA BROWN D., CHURCHER C.M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RA COLE S.T., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE; 96181548.
 RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,
 RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
 RA COLE S.T.;
 RT "An integrated map of the genome of the tubercle bacillus,
 RT Mycobacterium tuberculosis H37RV, and comparison with Mycobacterium
 RT lepreae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:3132-3137(1996).
 DR EMBL: AL021931; CAA17372.1;
 KW Hypothetical protein.
 SQ SEQUENCE 197 AA; 21376 MW; FABC6607 CRC32;
 Query Match 65.4%; Score 34; DB 2; Length 197;
 Best Local Similarity 71.4%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 3 DPTGHSY 9
 I | | | | |
 Db 46 DPTSHAY 52
 RESULT 11
 ID O69955 PRELIMINARY; PRT: 224 AA.
 AC O69955;
 DT 01-AUG-1998 (TEMBLrel. 07, Created)
 DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE HYPOTHETICAL 24.1 KD PROTEIN (FRAGMENT).
 GN SC3F9.15.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA SEEGER K.J., HARRIS D.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE; 97000351.
 RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,
 RA KINASHI H., HOPWOOD D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL023862; CAA19638.1;
 KW Hypothetical protein.
 FT NON_TER 224 224
 SQ SEQUENCE 224 AA; 24075 MW; 75A2A272 CRC32;

Query Match 65.4%; Score 34; DB 2; Length 224;
 Best Local Similarity 85.7%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 DPTGHSY 9
 DB 204 DPTGSSY 210

RESULT 12
 ID O87835 PRELIMINARY; PRT; 215 AA.
 AC O87835;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE HYPOTHETICAL 24.2 KD PROTEIN (FRAGMENT).
 GN SC8A6.01.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA SEGER K.J., HARRIS D.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE; 97000351.
 RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,
 RA KINASHI H., HOPWOOD D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL031013; CAA19773.1; -.
 KW Hypothetical protein.
 FT NON_TER 1
 FT NON_TER 215 215
 SQ SEQUENCE 215 AA; 24196 MW; F7643103 CRC32;

Query Match 65.4%; Score 34; DB 2; Length 215;
 Best Local Similarity 85.7%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 DPTGHSY 9
 DB 15 DPTGSSY 21

RESULT 13
 ID Q14969 PRELIMINARY; PRT; 268 AA.
 AC Q14969;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE T(3;5)(Q25.1:P34) FUSION GENE NPM-MLF1.
 GN NPM-MLF1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96152893.

RA YONEDA-KATO N., LOOK A.T., KIRSTEIN M.N., VALENTINE M.B.,
 RA RAIMONDI S.C., COHEN K.J., CARROLL A.J., MORRIS S.W.;
 RT "The t(3;5)(q25.1;q34) of myelodysplastic syndrome and acute myeloid
 leukemia produces a novel fusion gene, NPM-MLF1.";
 RL Oncogene 12:265-275(1996).
 DR EMBL; L49054; AAA99997.1; -.
 SQ SEQUENCE 268 AA; 30627 MW; F1B19980 CRC32;

Query Match 65.4%; Score 34; DB 4; Length 268;
 Best Local Similarity 71.4%; Pred. No. 39;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DPTGHSY 9
 DB 99 DPNHSHF 105

RESULT 14
 ID O75339 PRELIMINARY; PRT; 1184 AA.
 AC O75339;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE CARTILAGE INTERMEDIATE LAYER PROTEIN.
 GN CILP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ARTICULAR CARTILAGE;
 RX MEDLINE; 98389785.
 RA LORENZO P., NEAME P., SOMMARIN Y., HEINEGARD D.;
 RT "Cloning and deduced amino acid sequence of a novel cartilage protein
 (CILP) identifies a proform including a nucleotide
 J. Biol. Chem. 273:23469-23475(1998).
 RL EMBL; AF035408; AAC33838.1; -.
 DR PFAM; PF00047; ig; 1; 1.
 DR PFAM; PF00090; tsp_1; 1.
 SQ SEQUENCE 1184 AA; 132538 MW; 3E6F9774 CRC32;

Query Match 65.4%; Score 34; DB 4; Length 1184;
 Best Local Similarity 71.4%; Pred. No. 19e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DPTGHSY 9
 DB 1068 DPLGHNY 1074

RESULT 15
 ID O75862 PRELIMINARY; PRT; 347 AA.
 AC O75862;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE MAGE-B1.
 GN MAGE-B1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MUZYNY D., ARONSON A.D., ADAMS C., BRUNDAGE E., BUNAC C., CARVELLI K.,
 RA CHACKO J., CHEN J., DI W., DING Y., DUGAN S., DURBIN J., FORCUM J.,
 RA GANESH R., GARCIA C., GOODMAN M., GORRELL J.H., HAYWOOD M.,
 RA HERNANDEZ J., JACKSON L., JIN S., KAMPAL R., KARPATY S., KOVAR C.,
 RA LEAL B., LI Y., LICHTARGE O., LIU W., LOGAN O., LU J., LY T.,

RA MARTINEZ C., OSWAL G., PEREZ L., RASHID N.D., ROWLAND K., SAVAGE L.,
RA SCHERER S.E., SHEN H., SIMON M., STOVALL K., TIMMS K.M., TODD J.,
RA VO O., WILLIAMSON A., WORLEY K.C., YU W., CHINAULT C., NELSON D.,
RA GIBBS R.A.;
RT "Direct Submission";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005185; AAD10634.1; -;
DR PFAM; PF01454; MAGE; 1.
SQ SEQUENCE 347 AA; 39037 MW; 48A34904 CRC32;

Query Match 65.4%; Score 34; DB 4; Length 347;
Best Local Similarity 55.6%; Pred. No. 52;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EADPTGHSY 9
| : : : : :
Db 167 EDNPSGTY 175

Search completed: December 11, 1999, 01:16:39
Job time: 3819 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 1999, 02:36:51 ; Search time 2290.15 Seconds
(without alignments)
7879.429 Million cell updates/sec

Title: US-08-819-669D-8
Perfect score: 5674
Sequence: 1 CCCGGGACCACTGGCATC.....TAATGATCTCGGTGGATCC 5674

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database :

GenEmbl:

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_p11.*
- 8: gb_p12.*
- 9: gb_p13.*
- 10: gb_p14.*
- 11: gb_p15.*
- 12: gb_ro.*
- 13: gb_st.*
- 14: gb_sts.*
- 15: gb_sy.*
- 16: gb_un.*
- 17: gb_v1.*
- 18: em_fun.*
- 19: em_htg.*
- 20: em_hum1.*
- 21: em_hum2.*
- 22: em_in.*
- 23: em_om.*
- 24: em_or.*
- 25: em_ov.*
- 26: em_pat.*
- 27: em_ph.*
- 28: em_pl.*
- 29: em_ro.*
- 30: em_sts.*
- 31: em_sy.*
- 32: em_un.*
- 33: em_v1.*
- 34: gb_hcg1.*
- 35: gb_hcg2.*
- 36: gb_in1.*
- 37: gb_in2.*
- 38: em_ba1.*
- 39: em_ba2.*
- 40: em_hum3.*
- 41: em_hum4.*
- 42: gb_p14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5674	100.0	5674	5	I24013 Sequence 1

2	5674	100.0	5674	5	I36923
3	5532.6	97.5	156854	10	HSU082672
4	2777.2	48.9	49375	42	AF134576
5	2655	45.8	11495	9	HSU10687
6	2653.2	46.8	15898	11	U82696
7	2513.6	44.3	4895	9	HSU10688
8	2429.6	42.8	4736	9	HSU10690
9	2422.8	42.7	4741	9	HSU10689
10	2419	42.6	2419	5	I36922
11	2408	42.4	2420	5	AR007331
12	2408	42.4	2420	5	HUMAG1A
13	2307.6	40.7	54155	10	HSAF002994
14	2307.6	40.7	37781	11	AF030261
15	2306	40.6	48574	10	HSAF002996
16	2199.6	38.8	48574	10	HSAF002996
17	2117.8	37.3	4559	9	HUMAGE2X
18	2058.8	36.3	4157	5	I36924
19	1945	34.3	4523	9	HUMAGE12X
20	1762.6	31.1	4204	9	HSU03735
21	1678.8	29.6	2531	5	I36928
22	1678.8	29.6	2531	5	I36929
23	1677	29.6	3871	9	HSU10691
24	1556.6	27.4	4327	10	HSU69568
25	1536.4	27.1	3680	9	HSU10692
26	1504.6	26.5	2305	5	I36932
27	1410.8	24.9	2226	5	I36931
28	1353.8	23.9	73360	11	HSU66083
29	1247.4	22.0	3839	9	HSU10693
30	1226.6	21.6	2931	9	HSU10694
31	1051	18.5	3672	9	HSU10686
32	988.8	17.4	1640	5	I36926
33	900.2	15.9	1810	5	I36935
34	866	15.3	3510	9	HSU10685
35	830.6	14.6	1947	5	I36934
36	815.2	14.4	1412	5	I36936
37	760	13.4	1022	9	HUMAGEA
38	756.8	13.3	1022	9	HUMAGEC
39	750.8	13.2	1068	5	I36930
40	733.6	12.9	1111	14	G06157
41	701.6	12.4	954	9	HSU10340
42	695.8	12.3	1019	9	HUMAGEB
43	664.2	11.7	76133	10	HSU11148
44	629.8	11.1	945	9	HSU10339
45	583.4	10.3	1107	5	I36938

ALIGNMENTS

RESULT	1	I24013	5674 bp	DNA	PAT	21-NOV-1996
LOCUS	I24013	Sequence 1 from patent US 5541104.				
DEFINITION	I24013					
ACCESSION	91603883					
NID	I24013.1	GI:1603883				
VERSION						
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 5674)					
AUTHORS	Chen, Y., Stockert, E., Chen, Y., Garin-Chesa, P., Rettig, W.J., van der Bruggen, P., Boon-Falleur, T. and Old, L.J.					
TITLE	Monoclonal antibodies which bind to tumor rejection antigen precursor mage-1					
JOURNAL	Patent: US 5541104-A 1 30-JUL-1996;					
FEATURES	Location/Qualifiers					
SOURCE	1 .5674					
BASE COUNT	1276 a, 1644 c, 1569 g, 1185 t					
ORIGIN						

Query Match									
Best Local Similarity 100.0%; Score 5674; DB 5; Length 5674;									
Matches 5674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CCCCGGGACACAC	TGGCATCCCTCCCTACCA	CCCCCAATCCCTTACCC	CCCCACCC	60			
Db	1	CCCCGGGACACAC	TGGCATCCCTCCCTACCA	CCCCCAATCCCTTACCC	CCCCACCC	60			
Qy	61	ATCCAAACATCTT	CACGCTACACCCCAAGC	CAAGCAGAGATCCGG	TTCACCCCTG	120			
Db	61	ATCCAAACATCTT	CACGCTACACCCCAAGC	CAAGCAGAGATCCGG	TTCACCCCTG	120			
Qy	121	CTCTAACCCAGGA	AGCCAGGTGCCAGATGT	GAGCCACTGACTTGAG	CATTAGTGG	180			
Db	121	CTCTAACCCAGGA	AGCCAGGTGCCAGATGT	GAGCCACTGACTTGAG	CATTAGTGG	180			
Qy	181	TTAGAGAGACGAG	TTTTCGGTCTGAGGGCG	GTGAGATCGGTGGAG	GAAGGGG	240			
Db	181	TTAGAGAGACGAG	TTTTCGGTCTGAGGGCG	GTGAGATCGGTGGAG	GAAGGGG	240			
Qy	241	CCCAGCTCTGTAA	GAGGCAAGGTGACATG	CTGAGGGAGGACTGAG	GACCCACTTACCCC	300			
Db	241	CCCAGCTCTGTAA	GAGGCAAGGTGACATG	CTGAGGGAGGACTGAG	GACCCACTTACCCC	300			
Qy	301	AGATAGAGACCCC	AAATAATATCCCTT	CATGCCAGTCTGAG	CACTCTGGTGGACTTC	360			
Db	301	AGATAGAGACCCC	AAATAATATCCCTT	CATGCCAGTCTGAG	CACTCTGGTGGACTTC	360			
Qy	361	TCAGGCTGGGCA	CCCCCAGCCCTTGC	TGCTTAACCACTGGG	AGCTCAGAG	420			
Db	361	TCAGGCTGGGCA	CCCCCAGCCCTTGC	TGCTTAACCACTGGG	AGCTCAGAG	420			
Qy	421	CTCCGTGTGATC	AGGAGGCTGCTTAGG	AGGCGAGCGTCTG	CCAGACAT	480			
Db	421	CTCCGTGTGATC	AGGAGGCTGCTTAGG	AGGCGAGCGTCTG	CCAGACAT	480			
Qy	481	CATGCTCAGGAT	TTCAAGGAGGCTGAG	GGTCCCTAAGACCC	CACCTCCCGTGACCCAA	540			
Db	481	CATGCTCAGGAT	TTCAAGGAGGCTGAG	GGTCCCTAAGACCC	CACCTCCCGTGACCCAA	540			
Qy	541	CCCCACATCCCC	ACCCATCCCTCAAC	CCCTGATGCCATCCG	CCAGCCATCCACCC	600			
Db	541	CCCCACATCCCC	ACCCATCCCTGATGCC	ATCCGCCAGCCATCCAC	CCCTCCACCC	600			
Qy	601	CCCCACATCCCC	ACCCATCCCTCAAC	CCCTGATGCCATCCG	CCAGCCATCCACCC	660			
Db	601	CCCCACATCCCC	ACCCATCCCTCAAC	CCCTGATGCCATCCG	CCAGCCATCCACCC	660			
Qy	661	CACCCCAACCC	CAACCCAGCCACT	TCCACCCCAAGCAG	GATCCCGTCCCG	720			
Db	661	CACCCCAACCC	CAACCCAGCCACT	TCCACCCCAAGCAG	GATCCCGTCCCG	720			
Qy	721	CCAGGAACATCC	GGGTCCCGGATGTAG	CCCACTGACTTGC	ATGCGGACAGA	780			
Db	721	CCAGGAACATCC	GGGTCCCGGATGTAG	CCCACTGACTTGC	ATGCGGACAGA	780			
Qy	781	GAAGCAGGTTT	TCATCTGAGGACGCG	TAGAGTTGCGCC	CAAGAACTGACCC	840			
Db	781	GAAGCAGGTTT	TCATCTGAGGACGCG	TAGAGTTGCGCC	CAAGAACTGACCC	840			
Qy	841	CTCTGTGAGGA	AGCAAGGTGAGAGG	CTGAGGAGTGA	GAGACCCGCCACTCAA	900			
Db	841	CTCTGTGAGGA	AGCAAGGTGAGAGG	CTGAGGAGTGA	GAGACCCGCCACTCAA	900			
Qy	901	GAGAGCCCAAT	ATTCCAGCCCGCCCT	TGCTGCCAGCCCT	GCCACCCCGGGA	960			
Db	901	GAGAGCCCAAT	ATTCCAGCCCGCCCT	TGCTGCCAGCCCT	GCCACCCCGGGA	960			
Qy	961	CGTCTAGCCT	GGGCTGCCCGAC	CCCTGCTCCAA	AGCCTTGAGAGAC	1020			
Db	961	CGTCTAGCCT	GGGCTGCCCGAC	CCCTGCTCCAA	AGCCTTGAGAGAC	1020			
Qy	1021	TTCTCCCAAG	CTCTGGAATCAG	AGGTGCTGTG	TGACCGGCGGACTGG	1080			

Db	1021	TTCTCCCAAG	CTCTGGAATCAG	AGGTGCTGTG	TGACCGGCGGACTGG	1080			
Qy	1081	GCAGGACAGG	CTCTGCCAGG	CATCAAGATCAG	CAACCCAGAGGAGG	1140			
Db	1081	GCAGGACAGG	CTCTGCCAGG	CATCAAGATCAG	CAACCCAGAGGAGG	1140			
Qy	1141	CCAGAGCTC	AGCTCCCAATCC	CACTCCCACTCC	CAATTCGCATTC	1200			
Db	1141	CCAGAGCTC	AGCTCCCAATCC	CACTCCCACTCC	CAATTCGCATTC	1200			
Qy	1201	CCCATCTCT	CTCAGCTAC	ACCTCCACCC	CACTCTCTCTCT	1260			
Db	1201	CCCATCTCT	CTCAGCTAC	ACCTCCACCC	CACTCTCTCTCT	1260			
Qy	1261	ACCTCCAG	CCCCAGCAG	CCCAACCC	CTCTGCCCACCT	1320			
Db	1261	ACCTCCAG	CCCCAGCAG	CCCAACCC	CTCTGCCCACCT	1320			
Qy	1321	CCCACCTC	ATCTCTCT	CTCATGTG	CCCCCACTCC	1380			
Db	1321	CCCACCTC	ATCTCTCT	CTCATGTG	CCCCCACTCC	1380			
Qy	1381	GGTTTGCC	CTGCTCTCA	CCAGGAGG	CCCTGGT	1440			
Db	1381	GGTTTGCC	CTGCTCTCA	CCAGGAGG	CCCTGGT	1440			
Qy	1441	TGAACCTC	AGATCTGAG	AGAGCCAG	GTTCATTTA	1500			
Db	1441	TGAACCTC	AGATCTGAG	AGAGCCAG	GTTCATTTA	1500			
Qy	1501	ATCCACTG	AGGAGGAG	TGTTTAG	CTCTGTGAG	1560			
Db	1501	ATCCACTG	AGGAGGAG	TGTTTAG	CTCTGTGAG	1560			
Qy	1561	ACTGAGAG	GCACACAC	CCCCAGG	TAGATG	1620			
Db	1561	ACTGAGAG	GCACACAC	CCCCAGG	TAGATG	1620			
Qy	1621	CCAGCCCT	GACACCC	CGCCAGG	ACAGATGT	1680			
Db	1621	CCAGCCCT	GACACCC	CGCCAGG	ACAGATGT	1680			
Qy	1681	CAGTGCAC	TTAACCC	CACAGG	CAATCTGT	1740			
Db	1681	CAGTGCAC	TTAACCC	CACAGG	CAATCTGT	1740			
Qy	1741	GGTCAGAG	AGGAGG	CCCGG	CCAGGATCA	1800			
Db	1741	GGTCAGAG	AGGAGG	CCCGG	CCAGGATCA	1800			
Qy	1801	ACCTGGAG	GGGAGG	CTCAGG	GTTCCT	1860			
Db	1801	ACCTGGAG	GGGAGG	CTCAGG	GTTCCT	1860			
Qy	1861	CCACTC	ACATTC	CCCATAC	CTACCC	1920			
Db	1861	CCACTC	ACATTC	CCCATAC	CTACCC	1920			
Qy	1921	TCACCC	CAAGG	AGGAGG	CTGAGG	1980			
Db	1921	TCACCC	CAAGG	AGGAGG	CTGAGG	1980			
Qy	1981	GGGTCTG	ATGAGG	AGGCTTGA	ACAGGCTC	2040			
Db	1981	GGGTCTG	ATGAGG	AGGCTTGA	ACAGGCTC	2040			
Qy	2041	TGGGAGAT	GAGGAGG	CGCTCAG	AGGACCC	2100			
Db	2041	TGGGAGAT	GAGGAGG	CGCTCAG	AGGACCC	2100			
Qy	2101	ACTGAG	GCTGCC	ACTCA	AGATCAG	2160			
Db	2101	ACTGAG	GCTGCC	ACTCA	AGATCAG	2160			

QY 4321 CTCAGTCCCTTGAGTGGTCTTTTGGCAATGACGTGAAGGAACAGACACCCACCGGCCA 4380
DB 4321 CTCAGTCCCTTGAGTGGTCTTTTGGCAATGACGTGAAGGAACAGACACCCACCGGCCA 4380
QY 4381 CTCCTATGTCTCTGACCTGCTAGTCTCTCTATGATGGCTGCTGGGTGATATCA 4440
DB 4381 CTCCTATGTCTCTGACCTGCTAGTCTCTCTATGATGGCTGCTGGGTGATATCA 4440
QY 4441 GATCATGCCCAAGACAGCTTCCTGATAATTTGCTGCTGATGATGCAATGGAGGGCGG 4500
DB 4441 GATCATGCCCAAGACAGCTTCCTGATAATTTGCTGCTGATGATGCAATGGAGGGCGG 4500
QY 4501 CCATGCTCCTGAGGAGGAATCTGGGAGGAGCTGAGTGTGATGAGGTGATATGGGAG 4560
DB 4501 CCATGCTCCTGAGGAGGAATCTGGGAGGAGCTGAGTGTGATGAGGTGATATGGGAG 4560
QY 4561 GGAGCAGAGTCCCTATGGGAGGCCAGGAGCTGCTACCCCAAGATTTGGTGCAGGAAA 4620
DB 4561 GGAGCAGAGTCCCTATGGGAGGCCAGGAGCTGCTACCCCAAGATTTGGTGCAGGAAA 4620
QY 4621 GTACCTGGAGTACGCGAGGTCCGCGACAGTATCCGCGACGCTATGAGTTCCTGTGGGT 4680
DB 4621 GTACCTGGAGTACGCGAGGTCCGCGACAGTATCCGCGACGCTATGAGTTCCTGTGGGT 4680
QY 4681 CCAAGGGCCCTCGCTGAAACAGCTATGTGAAGTCTCTTGAGTGTGATCAAGGTCACT 4740
DB 4681 CCAAGGGCCCTCGCTGAAACAGCTATGTGAAGTCTCTTGAGTGTGATCAAGGTCACT 4740
QY 4741 GCAAGAGTTCGCTTTTCTCCCATCCCTGCGTGAAGCAGCTTTGAGAGGAGGAGAG 4800
DB 4741 GCAAGAGTTCGCTTTTCTCCCATCCCTGCGTGAAGCAGCTTTGAGAGGAGGAGAG 4800
QY 4801 GGAGTCTGAGCATGAGTGTGACCAAGCCAGTGGGAGGAGTGGCCAGTGCACCTT 4860
DB 4801 GGAGTCTGAGCATGAGTGTGACCAAGCCAGTGGGAGGAGTGGCCAGTGCACCTT 4860
QY 4861 CCAGGGCCGCTCCAGCAGCTTCCTCCCTGCTGCTGACATGAGGCCCATCTTCACTCT 4920
DB 4861 CCAGGGCCGCTCCAGCAGCTTCCTCCCTGCTGCTGACATGAGGCCCATCTTCACTCT 4920
QY 4921 GAAAGAGCGGTGAGTGTCTCAGTAGTAGTGTCTGCTTATTTGGGTGACTTGGAGATT 4980
DB 4921 GAAAGAGCGGTGAGTGTCTCAGTAGTAGTGTCTGCTTATTTGGGTGACTTGGAGATT 4980
QY 4981 TATCTTTGTTCTCTTTTGGAAATCTTCAATGTTTTTTTTTAAAGGATGTTGAATGAAC 5040
DB 4981 TATCTTTGTTCTCTTTTGGAAATCTTCAATGTTTTTTTTTAAAGGATGTTGAATGAAC 5040
QY 5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTCAACAGCTTCTGTGTATATAGTTTAAAGG 5100
DB 5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTCAACAGCTTCTGTGTATATAGTTTAAAGG 5100
QY 5101 TAAAGTCTGTGTTTTTATTCAGATTTGGGAATCCATCTATTTTGTGAATTTGGGATAAT 5160
DB 5101 TAAAGTCTGTGTTTTTATTCAGATTTGGGAATCCATCTATTTTGTGAATTTGGGATAAT 5160
QY 5161 AACAGCAGTGAATAAGTACTTGAATAATGTGAATAATGAGCAGTAAATAGATGAGATAA 5220
DB 5161 AACAGCAGTGAATAAGTACTTGAATAATGTGAATAATGAGCAGTAAATAGATGAGATAA 5220
QY 5221 AGAATCTAAAGAAATTAAGAGATAGTCAATTTCTGCTTATACCTTCAGTCTATTTCTGAAA 5280
DB 5221 AGAATCTAAAGAAATTAAGAGATAGTCAATTTCTGCTTATACCTTCAGTCTATTTCTGAAA 5280
QY 5281 ATTTTAAAGATATATGCATACCTTGGATTTCCTTGGCTTCTTTCAGATGTAAGAGAAAT 5340
DB 5281 ATTTTAAAGATATATGCATACCTTGGATTTCCTTGGCTTCTTTCAGATGTAAGAGAAAT 5340
QY 5341 TAAATCTGAATAAAGAAATTTCTGCTGTTCTGCTGCTCTTTTCTCTCCATGCACTGAGCA 5400
DB 5341 TAAATCTGAATAAAGAAATTTCTGCTGTTCTGCTGCTCTTTTCTCTCCATGCACTGAGCA 5400
QY 5401 TCTGCTTTTGGAGGGCCCTGGGTTAGTGTGAGATGCTTAAGGTAAGCCAGACTCATAC 5460

DB 5401 TCTGCTTTTGGAGGGCCCTGGGTTAGTGTGAGATGCTAAGGTAAGCCAGACTCATAC 5460
QY 5461 CCACCCATAGGTCGTAGAGTCTAGGAGCTGACAGTACAGTAATCGAGTGGCAAGATGTC 5520
DB 5461 CCACCCATAGGTCGTAGAGTCTAGGAGCTGACAGTACAGTAATCGAGTGGCAAGATGTC 5520
QY 5521 CTCTAAAGATGTAGGAAAAAGTGAAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580
DB 5521 CTCTAAAGATGTAGGAAAAAGTGAAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580
QY 5581 GTGAGTGTCAATGCCCTGAGCTGGGCAATTTGGGGCTTTGGGAACTGCAGTTCCTTCT 5640
DB 5581 GTGAGTGTCAATGCCCTGAGCTGGGCAATTTGGGGCTTTGGGAACTGCAGTTCCTTCT 5640
QY 5641 GGGGAGCTGATGTATGATCTTTGGGTGATCC 5674
DB 5641 GGGGAGCTGATGTATGATCTTTGGGTGATCC 5674

RESULT 2
LOCUS 136923 5674 bp DNA PAT 21-APR-1997
DEFINITION Sequence 8 from patent US 5612201.
ACCESSION 136923
NID 92084883
VERSION 136923.1 GI:2084883
KEYWORDS
SOURCE Unknown:
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5674)
AUTHORS De Plaen,E., Boon-Falleur,T., Lethe ,B., Szikora,J., De Smet,C. and Chomez,P.
TITLE Isolated nucleic acid molecules useful in determining expression of a tumor rejection antigen precursor
JOURNAL Patent: US 5612201-A 8 18-MAR-1997;
FEATURES Location/Qualifiers
source 1. 5674
BASE COUNT 1276 a, 1644 c 1569 g 1185 t
ORIGIN

Query Match 100.0%; Score 5674; DB 5; Length 5674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGGGGACACCTGGCATCCCTCCCTACACCCCAATCCCTCCCTTTACGCCACCC 60
DB 1 CCAGGGGACACCTGGCATCCCTCCCTACACCCCAATCCCTCCCTTTACGCCACCC 60
QY 61 ATCAAAACATCTTCACGCTCACCCCAAGCCAGCCAGGAGAGTAATCCGGTTCACCCCTG 120
DB 61 ATCAAAACATCTTCACGCTCACCCCAAGCCAGCCAGGAGAGTAATCCGGTTCACCCCTG 120
QY 121 CTCTCAACCCAGGAAGCCAGGTGCCAGATGTGACCCACCTGACCTGAGCATTTAGTGG 180
DB 121 CTCTCAACCCAGGAAGCCAGGTGCCAGATGTGACCCACCTGACCTGAGCATTTAGTGG 180
QY 181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGGGCTTTGAGATCGGTGGAGGAAGCGGG 240
DB 181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGGGCTTTGAGATCGGTGGAGGAAGCGGG 240
QY 241 CCAGCTCTGTAGAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300
DB 241 CCAGCTCTGTAGAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300
QY 301 AGATAGAGGACCCCAATAATCTCTTCATGCCAGTCTCGACCATCTGGTGGTGGACTTC 360
DB 301 AGATAGAGGACCCCAATAATCTCTTCATGCCAGTCTCGACCATCTGGTGGTGGACTTC 360
QY 361 TCAGGCTGGCCACCCCCAGCCCTTCTGCTTAAACACTGCGGAGCTCGAAGTCAGAG 420

Db	361	TCAGGCTGGGCCACCCCGCAGCCCCCTTGCTGTCTAAACCACTGGGACTCGAAGTCA	GAG 420
Qy	421	CTCCGTGTGATCAGGAAGGGTGTCTTAGGAGAGGGCAGCGTCCAGGCTCTGCAGACAT	480
Db	421	CTCCGTGTGATCAGGNAGGGTGTCTTAGGAGAGGGCAGCGTTCAGGCTCTGCCAGACAT	480
Qy	481	CATGCTCAGGATTCTCAAGAGAGGGTGAAGGTCCCTTAAGACCCCACTCCCCTGACCCAAC	540
Db	481	CATGCTCAGGATTCTCAAGAGAGGGTGAAGGTCCCTTAAGAGCCCACTCCCCTGACCCAAC	540
Qy	541	CCCCACTCCAATGCTCACCTCCCGTGACCACAACCCCTCTTTATTTGTCAATTCACACCCCA	600
Db	541	CCCCACTCCAATGCTCACCTCCCGTGACCACAACCCCTCTTTATTTGTCAATTCACACCCCA	600
Qy	601	CCCCACATCCCCCACCCCATCCCTCAACCCCTGATGCCATCGGCCAGGCCATTCACCCCT	660
Db	601	CCCCACATCCCCCACCCCATCCCTCAACCCCTGATGCCATCGGCCAGGCCATTCACCCCT	660
Qy	661	CACCCCAACCCCAACCCCAACCCCACTCCCAACCCCAACCCCAAGGATCCCGTTCCCG	720
Db	661	CACCCCAACCCCAACCCCAACCCCACTCCCAACCCCAACCCCAAGGATCCCGTTCCCG	720
Qy	721	CCAGGAAACATCCGGTGC CGGGATGTGAGCCACTGACTTCGGCATTTGTGGGCACAGA	780
Db	721	CCAGGAAACATCCGGTGC CGGGATGTGAGCCACTGACTTCGGCATTTGTGGGCACAGA	780
Qy	781	GAAGCAGGTTTCCATTCTGAGGAGCGGTAGAGTTCGGCCGAAGGAACCTGACCCAGG	840
Db	781	GAAGCAGGTTTCCATTCTGAGGAGCGGTAGAGTTCGGCCGAAGGAACCTGACCCAGG	840
Qy	841	CTCTGTGAGGAGCAAGGTGAGAGGTGAGGAGGACTGAGSACCCCGCCACTCAAATA	900
Db	841	CTCTGTGAGGAGCAAGGTGAGAGGTGAGGAGGACTGAGSACCCCGCCACTCAAATA	900
Qy	901	GAGAGCCCCAAATATTCCAGCCCCCGCTTGCTGCCAGCCCTGGCCACCCCGGGGAAGA	960
Db	901	GAGAGCCCCAAATATTCCAGCCCCCGCTTGCTGCCAGCCCTGGCCACCCCGGGGAAGA	960
Qy	961	CGTCTCAGCCTGGGCTGCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAAGTTC	1020
Db	961	CGTCTCAGCCTGGGCTGCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAAGTTC	1020
Qy	1021	TTCTCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGCGAGSACTGGTTAGGAGAGG	1080
Db	1021	TTCTCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGCGAGSACTGGTTAGGAGAGG	1080
Qy	1081	GCAGGGCACAGGCTCTGC CAGGCATCAAGATCAGCACCCACAGAGGAGGGCTGTGGGCC	1140
Db	1081	GCAGGGCACAGGCTCTGC CAGGCATCAAGATCAGCACCCACAGAGGAGGGCTGTGGGCC	1140
Qy	1141	CCAAGACTGCAC TCCAATCCCACCTCCACCCATTGCGATTCCTCCATCCCACCCCAACC	1200
Db	1141	CCAAGACTGCAC TCCAATCCCACCTCCACCCATTGCGATTCCTCCATCCCACCCCAACC	1200
Qy	1201	CCCATCTCCTCAGCTACACTCCACCCCATCCCTACTCTCTACTCCGTCAGCTGACCAAC	1260
Db	1201	CCCATCTCCTCAGCTACACTCCACCCCATCCCTACTCTCTACTCCGTCAGCTGACCAAC	1260
Qy	1261	ACCTTCAGCCCCCAGCACAGCCCCAACCCCTTCTGCGACCTTCACCTGACCTGCCCAAC	1320
Db	1261	ACCTTCAGCCCCCAGCACAGCCCCAACCCCTTCTGCGACCTTCACCTGACCTGCCCAAC	1320
Qy	1321	CCGACCTCATCTCTCTCATGTGCCCACTCCCATCGCTCCCCCATCTGCGAGAAATCC	1380
Db	1321	CCGACCTCATCTCTCTCATGTGCCCACTCCCATCGCTCCCCCATCTGCGAGAAATCC	1380
Qy	1381	GGTTTGCCCTCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGCATGTGAAACCACTGACT	1440
Db	1381	GGTTTGCCCTCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGCATGTGAAACCACTGACT	1440
Qy	1441	TGAACCTCACAGATCTGAGAGAAGCAGGTTCAATTAATGTTCTGAGGGCGGCTTGAG	1500
Db	1441	TGAACCTCACAGATCTGAGAGAAGCAGGTTCAATTAATGTTCTGAGGGCGGCTTGAG	1500

QY 2581 ACTATGTGAGGAATTTGGGGTTGAGGAAGCACAGCGCTGGCAGGAATAAGATGAGT 2640
DB 2581 ACTATGTGAGGAATTTGGGGTTGAGGAAGCACAGCGCTGGCAGGAATAAGATGAGT 2640
QY 2641 GAGACAGACAAGGCTATTGGAATTCACACCCCGAGAACCAAGGGGTGAGCCTTGACACG 2700
DB 2641 GAGACAGACAAGGCTATTGGAATTCACACCCCGAGAACCAAGGGGTGAGCCTTGACACG 2700
QY 2701 TCACCCAGGATGTGGCTTCTTTTTCACCTCTCTTCCAGATCTGGGGAGGTGAGGACCT 2760
DB 2701 TCACCCAGGATGTGGCTTCTTTTTCACCTCTCTTCCAGATCTGGGGAGGTGAGGACCT 2760
QY 2761 CATTTCTCAGAGGCTGACTCAGGTCAACGTAGGAGACCCCATCTGGTCTAAAGACAGAGGG 2820
DB 2761 CATTTCTCAGAGGCTGACTCAGGTCAACGTAGGAGACCCCATCTGGTCTAAAGACAGAGGG 2820
QY 2821 GTCCAGGATCTGCCATGCGTTTGGGTGAGGAACATGAGGAGGACTGAGGGTACCCAG 2880
DB 2821 GTCCAGGATCTGCCATGCGTTTGGGTGAGGAACATGAGGAGGACTGAGGGTACCCAG 2880
QY 2881 GACCAAGACACTGAGGAGACTGCACAGAAATCAGCCCTGCCCTGCTGCACCCAGAG 2940
DB 2881 GACCAAGACACTGAGGAGACTGCACAGAAATCAGCCCTGCCCTGCTGCACCCAGAG 2940
QY 2941 AGCATGGGCTGGGCGCTCTGCGAGGCTCTCCGTTATCTGGGATCAATTGATGTCAGGG 3000
DB 2941 AGCATGGGCTGGGCGCTCTGCGAGGCTCTCCGTTATCTGGGATCAATTGATGTCAGGG 3000
QY 3001 ACGGGAGGCTTGGTCTGAGAGGCTGGCTCAGGTCAAGTAGAGGAGCGTCCAGGCG 3060
DB 3001 ACGGGAGGCTTGGTCTGAGAGGCTGGCTCAGGTCAAGTAGAGGAGCGTCCAGGCG 3060
QY 3061 CTGCGAGGACTCAAGTGTAGGAGCAAGCGGCGACCTCACCCAGGACACATTAATTCACAT 3120
DB 3061 CTGCGAGGACTCAAGTGTAGGAGCAAGCGGCGACCTCACCCAGGACACATTAATTCACAT 3120
QY 3121 GAATTTGATATCTCTGTGCTTCCCTTCCCAAGGACCTAGGACGCTGTGGCCAGATGTTT 3180
DB 3121 GAATTTGATATCTCTGTGCTTCCCTTCCCAAGGACCTAGGACGCTGTGGCCAGATGTTT 3180
QY 3181 GTCCCTCTGCTTCCATCTTATCATGATGTGAATCTTGAATTTGGATTTCTCAG 3240
DB 3181 GTCCCTCTGCTTCCATCTTATCATGATGTGAATCTTGAATTTGGATTTCTCAG 3240
QY 3241 ACCAGCAAGGCGAGGATCCAGGCGCTCCAGGAAATAAAGGGCCCTCGCTGAGAA 3300
DB 3241 ACCAGCAAGGCGAGGATCCAGGCGCTCCAGGAAATAAAGGGCCCTCGCTGAGAA 3300
QY 3301 CAGAGGGGTCTATCCACTGCTGATGAGAGTGGGATGTACAGAGTCCAGCCACCCCTCTG 3360
DB 3301 CAGAGGGGTCTATCCACTGCTGATGAGAGTGGGATGTACAGAGTCCAGCCACCCCTCTG 3360
QY 3361 GTAGCACTGAGAGCCAGGGCTGTGCTGCGGTCTGACCCCTGAGGCGCCGTGGATTCCT 3420
DB 3361 GTAGCACTGAGAGCCAGGGCTGTGCTGCGGTCTGACCCCTGAGGCGCCGTGGATTCCT 3420
QY 3421 CTTCTGAGCTCCAGGAACAGGAGTGGGCTTGGTCTGAGACAGTATCTCAGGTC 3480
DB 3421 CTTCTGAGCTCCAGGAACAGGAGTGGGCTTGGTCTGAGACAGTATCTCAGGTC 3480
QY 3481 ACAGAGCAGAGGATGCACAGGCTGTCAGCAGTGAATGTTTGGCCCTGAATGCACACAA 3540
DB 3481 ACAGAGCAGAGGATGCACAGGCTGTCAGCAGTGAATGTTTGGCCCTGAATGCACACAA 3540
QY 3541 GGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGGCTCACCTCCCTACTG 3600
DB 3541 GGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGGCTCACCTCCCTACTG 3600
QY 3601 TCAGTCTCTAGAACTGACCTCTGCTGGCGGCTGTACCTCAGTACCTCTCACTTCT 3660
DB 3601 TCAGTCTCTAGAACTGACCTCTGCTGGCGGCTGTACCTCAGTACCTCTCACTTCT 3660

QY 3661 COTTGAGGTTTTCAGGAGCAGGCCAACCCAGAGACAGGATTCCTCTGAGGCCACAGAG 3720
DB 3661 COTTGAGGTTTTCAGGAGCAGGCCAACCCAGAGACAGGATTCCTCTGAGGCCACAGAG 3720
QY 3721 GAGCACCAGGAGAGATCTGTAAGTAGCCCTTTGTAGAGTCTCAAGGTTTCAAGTTCTC 3780
DB 3721 GAGCACCAGGAGAGATCTGTAAGTAGCCCTTTGTAGAGTCTCAAGGTTTCAAGTTCTC 3780
QY 3781 AGCTGAGGCTCTCACACACTCCCTCTCTCCCGAGGCTGTGGTCTTCATTGCCAGCT 3840
DB 3781 AGCTGAGGCTCTCACACACTCCCTCTCTCCCGAGGCTGTGGTCTTCATTGCCAGCT 3840
QY 3841 COTGCCACACTCTCTGCTGCTGCCCTGACGAGAGTCAATCATGTCTCTTTGAGCAGAGAG 3900
DB 3841 COTGCCACACTCTCTGCTGCTGCCCTGACGAGAGTCAATCATGTCTCTTTGAGCAGAGAG 3900
QY 3901 TCTGACTGCAAGCCTGAGGAAGCCCTTTGAGGCCCAACAGAGGCCCTGGGCTGTGTG 3960
DB 3901 TCTGACTGCAAGCCTGAGGAAGCCCTTTGAGGCCCAACAGAGGCCCTGGGCTGTGTG 3960
QY 3961 TGTGAGGCTGCACT 4020
DB 3961 TGTGAGGCTGCACT 4020
QY 4021 CACTGCTGGTCAACAGATCT 4080
DB 4021 CACTGCTGGTCAACAGATCT 4080
QY 4081 CATCAACTTCACTCAGAGAGGCAACCCAGTGGGTTTCCAGCAGCCGTTGAAGAGAGGG 4140
DB 4081 CATCAACTTCACTCAGAGAGGCAACCCAGTGGGTTTCCAGCAGCCGTTGAAGAGAGGG 4140
QY 4141 GCAAGCAGCT 4200
DB 4141 GCAAGCAGCT 4200
QY 4201 TGATTTGGTGGTTTCT 4260
DB 4201 TGATTTGGTGGTTTCT 4260
QY 4261 AATGCTGAGAGTGTCAATCAAAATTAAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4320
DB 4261 AATGCTGAGAGTGTCAATCAAAATTAAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4320
QY 4321 CTCTGAGTCTCTGAGCTGGTCTTTGGCATTGACCTGAAGAGCAGACCCACCGGCA 4380
DB 4321 CTCTGAGTCTCTGAGCTGGTCTTTGGCATTGACCTGAAGAGCAGACCCACCGGCA 4380
QY 4381 CTCCTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4440
DB 4381 CTCCTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4440
QY 4441 GATCATGCCCCAGACAGGCTTCTCTGATTAATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4500
DB 4441 GATCATGCCCCAGACAGGCTTCTCTGATTAATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4500
QY 4501 CCATGCTCTCTGAGGAGGAAATCTGGAGAGCTGAGTGTGATGAGGAGTGTATGATGGGAG 4560
DB 4501 CCATGCTCTCTGAGGAGGAAATCTGGAGAGCTGAGTGTGATGAGGAGTGTATGATGGGAG 4560
QY 4561 GGAGCAGAGTGCCTATGGGAGGCCAGGAAGCTGCTACCCCAAGATTTGGTGCAGGAAAA 4620
DB 4561 GGAGCAGAGTGCCTATGGGAGGCCAGGAAGCTGCTACCCCAAGATTTGGTGCAGGAAAA 4620
QY 4621 GTACCTGAGTACGCGAGGTCGCCGACAGTGTATCCCGACGCTATGAGTTCCTGTGGGT 4680
DB 4621 GTACCTGAGTACGCGAGGTCGCCGACAGTGTATCCCGACGCTATGAGTTCCTGTGGGT 4680
QY 4681 CCAAGGGCCCTCGCTGAAACAGCTATGTGAAGTCTTGAAGTGTGATGATCAAGGTCACT 4740
DB 4681 CCAAGGGCCCTCGCTGAAACAGCTATGTGAAGTCTTGAAGTGTGATGATCAAGGTCACT 4740
QY 4741 GCAAGAGTTCGCTTTTCT 4800

```
|||||
Db 4741 GCAGAGTTCCTTTCTTCCATCCCTCGGTGAAGCAGCTTTGAGAGAGGAGAG 4800
Qy 4801 GGAGTCTGAGCATGAGTTGACCAAGCCAGTGGAGGGGAGTGGCCAGTGCACCTT 4860
Db 4801 GGAGTCTGAGCATGAGTTGACCAAGCCAGTGGAGGGGAGTGGCCAGTGCACCTT 4860
Qy 4861 CCAGGGCGGCTCCAGCAGCTTCCCTCGTGTGACATGAGGCCCATCTTTCACCT 4920
Db 4861 CCAGGGCGGCTCCAGCAGCTTCCCTCGTGTGACATGAGGCCCATCTTTCACCT 4920
Qy 4921 GAAGAGAGCGTCAAGTCTTCAGTAGTAGTCTTCTGTTCTTATGGGTGACATGAGATT 4980
Db 4921 GAAGAGAGCGTCAAGTCTTCAGTAGTAGTCTTCTGTTCTTATGGGTGACATGAGATT 4980
Qy 4981 TATCTTTGTTCTTTTGGAAATGTTCAAAATGTTTTTTTAAAGGATGTTGAATGAAC 5040
Db 4981 TATCTTTGTTCTTTTGGAAATGTTCAAAATGTTTTTTTAAAGGATGTTGAATGAAC 5040
Qy 5041 TTCAGCATCCAAAGTTTATGAATGACAGCAGTCACACAGTCTCTGTGTATATAGTTAAGG 5100
Db 5041 TTCAGCATCCAAAGTTTATGAATGACAGCAGTCACACAGTCTCTGTGTATATAGTTAAGG 5100
Qy 5101 TAAGAGTCTGTGTTTATTCAGATGGGAATCCATCTCTATTTTGTGAATGGGATAAT 5160
Db 5101 TAAGAGTCTGTGTTTATTCAGATGGGAATCCATCTCTATTTTGTGAATGGGATAAT 5160
Qy 5161 AACAGCAGTGAATAAGTACTTGAATAATGAGAAATGAGCAGTAAATAGATGAGATAA 5220
Db 5161 AACAGCAGTGAATAAGTACTTGAATAATGAGAAATGAGCAGTAAATAGATGAGATAA 5220
Qy 5221 AGAACTAAAGAAATTAAGAGATAGTCAATTTCTGCTTATACCTCAGTCTATTCTGTAAA 5280
Db 5221 AGAACTAAAGAAATTAAGAGATAGTCAATTTCTGCTTATACCTCAGTCTATTCTGTAAA 5280
Qy 5281 ATTTTAAAGATATATGCATACCTGGATTTCTGCTGGCTCTTTTGGAGATGTAAGAGAAAT 5340
Db 5281 ATTTTAAAGATATATGCATACCTGGATTTCTGCTGGCTCTTTTGGAGATGTAAGAGAAAT 5340
Qy 5341 TAAATCTGAATAAGAAATTTCTCTGTTCACTGGCTCTTTTCTTCTCCATCAGTGCAGCA 5400
Db 5341 TAAATCTGAATAAGAAATTTCTCTGTTCACTGGCTCTTTTCTTCTCCATCAGTGCAGCA 5400
Qy 5401 TCTGCTTTTGGAAAGCCCTCGGTTAGTAGTGAGATGCTAAGGTAAGCCAGACTCATAC 5460
Db 5401 TCTGCTTTTGGAAAGCCCTCGGTTAGTAGTGAGATGCTAAGGTAAGCCAGACTCATAC 5460
Qy 5461 CCACCCATAGGTCGTAGAGTCTAGGAGCTGAGTCAAGTCAAGTGGCAGATGTC 5520
Db 5461 CCACCCATAGGTCGTAGAGTCTAGGAGCTGAGTCAAGTCAAGTGGCAGATGTC 5520
Qy 5521 CTCTAAAGATCTAGGAAAGCTGAGAGAGGCTGAGGCTGAGGCTCCGGTGAGAGTG 5580
Db 5521 CTCTAAAGATCTAGGAAAGCTGAGAGAGGCTGAGGCTGAGGCTCCGGTGAGAGTG 5580
Qy 5581 GTGAGTGTCAATCCCTGAGTGGGCAATTTTGGGCTTTTGGGAACTGCAAGTCCCTTCT 5640
Db 5581 GTGAGTGTCAATCCCTGAGTGGGCAATTTTGGGCTTTTGGGAACTGCAAGTCCCTTCT 5640
Qy 5641 GGGGAGCTGATGTAATGATCTTGGGTGGATCC 5674
Db 5641 GGGGAGCTGATGTAATGATCTTGGGTGGATCC 5674
```

```
RESULT 3
LOCUS HSU82672/c 156854 bp DNA PRI 13-MAY-1997
DEFINITION Human chromosome X clone Qc15B1, complete sequence.
ACCESSION U82672
NID 92078526
VERSION U82672.1 GI:2078526
KEYWORDS HTG.
SOURCE human.
```

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 156854)
AUTHORS Gloeckner, G., Rump, A., Nordstiek, G., Hinzmann, B., Kioschis, P.,
Heiss, N., Poustka, A., Bauer, D., Drescher, B., Knob, A. and
Rosenthal, A.
TITLE Sequencing and analysis of a region in Xq28 containing MAGE-1 and a
putative Zinc Finger Protein
JOURNAL Unpublished
AUTHORS Gernot, G.
REFERENCE
2 (bases 1 to 156854)
JOURNAL Direct Submission
TITLE Submitted (19-DEC-1996) Genome Analysis, IMB, Beutenbergstr. 11,
Jena 07745, Germany
FEATURES
Location/Qualifiers
1..156854
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xq28"
1..40357
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Qc15B1"
1057..1295
/repeat_region
2504..2666
/repeat_region
/rpt_family="Alu-Sz"
1295..15191
/mRNA
/organism="Homo sapiens"
/product="MAGE-1 Protein"
/clone="MAGE1"
1057..1295
/gene
/organism="Homo sapiens"
/product="MAGE-1 Protein"
/clone="MAGE1"
1057..1295
/CDS
/organism="Homo sapiens"
/product="MAGE-1 Protein"
/clone="MAGE1"
1057..1295
/codon_start=1
/organism="Homo sapiens"
/product="MAGE-1 Protein"
/clone="MAGE1"
1057..1295
/protein_id="AAB54061.1"
/db_xref="PIR:G2078527"
/db_xref="GI:2078527"
/translation="MSLEQSLHCKPEEALEAQEALGLVCVQAAASSSPLVLGTLE
EPTAGSTDPQSPQAGASPTTINFTQRPQEGSSREEGSTSCILLESFRAVI
TKVADLVGFLLLKRYAREPVTAKMLESVKNYKHFIFKAGSLSQLQVFDIVK
EADPTGHSYVLVTCGLSYDGLDNOIMPKTGTFLIIVLVIAHMEGHAPPEETWEEL
SVMEVDGRHSAYGEPKLLTODLVQEKYLEVRQVDSDPARVEFLWGPALAEYSY
VKVLEVIKVSARVRFPPSLRAALREEEGV"
9623..9892
/repeat_region
/rpt_family="Alu-Sz"
10698..10978
/repeat_region
/complement(10698..10978)
/rpt_family="Alu-Sx"
12218..12345
/repeat_region
/complement(12218..12345)
/rpt_family="MER42C"
12636..12851
/repeat_region
/complement(12636..12851)
/rpt_family="MER42A"
12909..13060
/repeat_region
/complement(12909..13060)
/rpt_family="MER42A"
13455..13825
/repeat_region
/complement(13455..13825)
/rpt_family="MLT1A"
13557..153516
/source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Qc15B1"
13836..13904
/repeat_region
/complement(13836..13904)
/rpt_family="LIMC2"
13939..14597
/repeat_region
/complement(13939..14597)
/rpt_family="LIMD1"
14609..14916
/repeat_region
/complement(14609..14916)
/rpt_family="L1"
15018..15180
/repeat_region
/complement(15018..15180)
/rpt_family="L1"
15186..15590
/repeat_region
/complement(15186..15590)
/rpt_family="MLT1B"
```

Query Match	97.5%	Score 5532.6;	DB 10;	Length 156854;
Best Local Similarity	99.3%;	pred. No. 0;		
Matches 5640;	Conservative 0;	Mismatches 29;	Indels 10;	Gaps 8
Qy 1	CCCGGGGCACTGGGATCCCTCCCTACCAACCCCAATCCCTCCCTTTACGGCACCC	60		
Db 9070	CCCGGGGACCACCTGGCATCCCTCCCTACCAACCCCAATCCCTCCCTTTACGGCACCC	9011		
Qy 61	ATCCAAACATCTTCACGCTCACCCCGCCCAAGCCAGCGAGAATCCGGTTCACCCCTG	120		
Db 9010	ATCCAAACATCTTCACGCTCACCCCGCCCAAGCCAGCGAGAATCCGGTTCACCCCTG	8951		
Qy 121	CTCTCAACCCAGGGAAGCCCAAGTGCCAGATGTGACGCCACTGACTTGAGCATTAGTGG	180		
Db 8950	CTCTCAACCCAGGGAAGCCCAAGTGCCAGATGTGACGCCACTGACTTGAGCATTAGTGG	8891		
Qy 181	TTAGAGAAACGAGGTTTTTCGGTCTGAGGGCGCGCTTGAGATCCGTTGAGGGAACGCGG	240		
Db 8890	TTAGAGAAACGAGGTTTTTCGGTCTGAGGGCGCGCTTGAGATCCGTTGAGGGAACGCGG	8831		
Qy 241	CCCA-GCTCTCTAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACCTATCCC	299		
Db 8830	CCCAGGCTCTGTAAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACCTATCCC	8771		
Qy 300	CAGATAGAGGACCCCAATATATCCCTTCATGCCAGTCTCTGGACCATCTGGTGGTGACTT	359		

|||||
Db 8770 CAGATAGAGACCCAAATAATCCCTTCATGCCAGTCCCTGGACATCTGGTGTGGACTT 8711
QY 360 CTGAGGCTGGCCACCCAGCCCTTGTCTGCTTAACACACTGGGACTCGGAAGTCAGA 419
Db 8710 CTCAGGCTGGCCACCCAGCCCTTGTCTGCTTAACACACTGGGACTCTGAAGTCAG 8651
QY 420 GCTCCGTGTGATCAGGAAGGGTGTGTAGGAGAGGCGCTCCAGGCTCTGCCAGACA 479
Db 8650 AGCTCGTGTGATCAGGAAGGGTGTGTAGGAGAGGCGCTCCAGGCTCTGCCAGACA 8591
QY 480 TCATGCTCAGGATCTCAAGGAGGGCTGAGGGTCCCTTAAGACCCCACTCCCGTGACCCAA 539
Db 8590 TCATGCTCAGGATCTCAAGGAGGGGTGAGGGTCCCTTAAGACCCCACTCCCGTGACCCAA 8531
QY 540 CCCCCACTCCAATGCTCACTCCCGTGACCCAGCCCTCTTCATTTGTCATTTCCAAACCCCC 599
Db 8530 CCCCCACTCCAATGCTCACTCCCGTGACCCAGCCCTCTTCATTTGTCATTTCCAAACCCCC 8471
QY 600 ACCCCACATCCCCACCCCAATCCCTCAACCCCTGATGCCCATCCGCCAGGCAATTCACCC 659
Db 8470 ACCCCACATCCCCACCCCAATCCCTCAACCCCTGATGCCCATCCGCCAGGCAATTCACCC 8411
QY 660 TCACCCCCACCCCAATCCCTCAACCCCTGATGCCCATCCGCCAGGCAATTCACCC 719
Db 8410 TCACCCCCACCCCAATCCCTCAACCCCTGATGCCCATCCGCCAGGCAATTCACCC 8351
QY 720 GCCAGGAACATCCGGTGGCCGGTGTGACGCACTGACTTGCATTTGCGGAGGAG 779
Db 8350 GCCAGGAACATCCGGTGGCCGGTGTGACGCACTGACTTGCATTTGCGGAGGAG 8291
QY 780 AGAAGCGAGGTTTCCATTTCTGAGGAGCGGCTAGAGTTGCGGCGAAGAACCTTGACCCAG 839
Db 8290 AGAAGCGAGGTTTCCATTTCTGAGGAGCGGCTAGAGTTGCGGCGAAGAACCTTGACCCAG 8231
QY 840 GCTCTGTGAGGAGCGAAGGTGAGAGGCTGAGGAGGACTGAGACCCCGCCACTCCAAT 899
Db 8230 GCTCTGTGAGGAGCGAAGGTGAGAGGCTGAGGAGGACTGAGACCCCGCCACTCCAAT 8171
QY 900 AGAGAGCCCAATATTTCCAGCCCGCCCTTGTCTGCGAGCCCTGGCCACCCCGGGAAG 959
Db 8170 AGAGAGCCCAATATTTCCAGCCCGCCCTTGTCTGCGAGCCCTGGCCACCCCGGGAAG 8111
QY 960 ACCTCTCAGCCTGGGTGCCCCAGACCCCTGCTTCCAAAAGCCTTGAGAGACACACAGTT 1019
Db 8110 ACCTCTCAGCCTGGGTGCCCCAGACCCCTGCTTCCAAAAGCCTTGAGAGACACACAGTT 8051
QY 1020 CTTCTCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCGGGGAGGACTGGTTAGGAGAG 1079
Db 8050 CTTCTCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCGGGGAGGACTGGTTAGGAGAG 7991
QY 1080 GGCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGACCCCAAGAGGGGAGGCTGTGGCC 1139
Db 7990 GGCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGACCCCAAGAGGGGAGGCTGTGGCC 7931
QY 1140 CCAAGAGCTCAGCTCCAATCCCACTCCCACTCCCACTTCGATTTCCCACTCCCACTCCCA 1199
Db 7930 CCAAGAGCTCAGCTCCAATCCCACTCCCACTTCGATTTCCCACTCCCACTCCCACTCCCA 7871
QY 1200 CCCCCTCTCAGCTACAGCTCCACCCCACTCCCACTTCGATTTCCCACTCCCACTCCCA 1259
Db 7870 CCCCCTCTCAGCTACAGCTCCACCCCACTCCCACTTCGATTTCCCACTCCCACTCCCA 7811
QY 1260 CACCTCCAGCCCGCAGCAGCCCAACCTTCTGCCACTTCACCTCAGCTCCAGCCCA 1319
Db 7810 CACCTCCAGCCCGCAGCAGCCCAACCTTCTGCCACTTCACCTCAGCTCCAGCCCA 7751
QY 1320 CCCCACCTCATCTCTCATGTGCCCCCACTCCCACTTCGATTTCCCACTCCCACTCCCA 1379
Db 7750 CCCCACCTCATCTCTCATGTGCCCCCACTCCCACTTCGATTTCCCACTCCCACTCCCA 7691
QY 1380 CGGTTGCCCCCTGCTCTCAACCCAGGGAAGCCCTGTGAGCCCGATGTGAACCACTGAC 1439
|||||

Db 7690 CGGTTTGCCTGCTCTCAACCCAGGGAAGCCCTGGATGGCCCGATGTGAACCACTGAC 7631
QY 1440 TTGAACCTCACAGATCTGAGAGAAGCCAGGTTTCAATTAATGTTCTGAGGGCGGCTTGA 1499
Db 7630 TTGAACCTCACAGATCTGAGAGAAGCCAGGTTTCAATTAATGTTCTGAGGGCGGCTTGA 7571
QY 1500 GATCCACTGAGGGAGTGGTTTGTAGGCTCTGTGAGGAGCAAGGTGAGATGCTGAGGGAG 1559
Db 7570 GATCCACTGAGGGAGTGGTTTGTAGGCTCTGTGAGGAGCAAGGTGAGATGCTGAGGGAG 7511
QY 1560 GACTGAGGAGGACACACCCAGGTAGATGGCCCAAAATGATCCAGTACACCCCTGCT 1619
Db 7510 GACTGAGGAGGACACACCCAGGTAGATGGCCCAAAATGATCCAGTACACCCCTGCT 7451
QY 1620 GCCAGCCCTGGACACCCCGCCAGGACAGATGCTCAGCTGGACACCCCGCTCCCGTC 1679
Db 7450 GCCAGCCCTGGACACCCCGCCAGGACAGATGCTCAGCTGGACACCCCGCTCCCGTC 7391
QY 1680 CCACTGCCACTTAACCCACAGGCAATCTGTAGTCAATAGCTTATGACCGGGGAGGTT 1739
Db 7390 CCACTGCCACTTAACCCACAGGCAATCTGTAGTCAATAGCTTATGACCGGGGAGGTT 7331
QY 1740 TGCTCAGGAGA-GGCAGGGCCAGGCAATCAAGGTCCA-GCATCCGCCCGCATTAGGTC 1797
Db 7330 TGCTCAGGAGAAGGCGCCAGGCAATCAAGGTCCAAGGTCCAGGCAATCCCGCGCATTAGGTC 7271
QY 1798 AGGACCTGGGAGGAACTGAGGTTCCCAACCCACACCTGCTCTCTCATCTCCACGCC 1857
Db 7270 AGGACCTGGGAGGAACTGAGGTTCCCAACCCACACCTGCTCTCTCATCTCCACGCC 7211
QY 1858 ACCCCACTCACATTTCCATACCTTACCCCTTACCCCACTCATCTGTCAGAACTCCCTG 1917
Db 7210 ACCCCACTCACATTTCCATACCTTACCCCTTACCCCACTCATCTGTCAGAACTCCCTG 7151
QY 1918 CTGTCACCCACGGAAGCCAGGCAATGGCGCCAGGCACTCGGATCTTACGCTCCCAT 1977
Db 7150 CTGTCACCCACGGAAGCCAGGCAATGGCGCCAGGCACTCGGATCTTACGCTCCCAT 7091
QY 1978 CAGGCTCTGATGAGGGAAGGGCTTGAACAGGGCTCAGGGGAGCAGAGGAGGGGCC 2037
Db 7090 CAGGG-CTGATGAGGGAAGGGCTTGAACAGGGCTCAGGGGAGCAGAGGAGGGGCC 7032
QY 2038 TACTCGAGATGAGGAGGCTCAGAGGACCCAGACCTTAGACACCCAGCCCTGTCT 2097
Db 7031 TACTCGAGATGAGGAGGCTCAGAGGACCCAGACCTTAGACACCCAGCCCTGTCT 6972
QY 2098 GAGACTGAGGCTGCCACTTCTGSCCTCAAGAATCAGAACGATGGGACACTCAGATT-GCAT 2156
Db 6971 GAGACTGAGGCTGCCACTTCTGSCCTCAAGAATCAGAACGATGGGACACTCAGATTAGCAT 6912
QY 2157 GGGGTTGGACCCAGGCTTGAAGGCTTACGCGGAGGAGGAGGAGGAGGAGGAGGAG 2216
Db 6911 AGGGTTGGACCCAGGCTTGAAGGCTTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 6852
QY 2217 CTTGGAATCCAGATCAGTGTGGACCTCGCCCTGAGAGGTCCAGGGCAGCGTGGCCACA 2276
Db 6851 CTTGGAATCCAGATCAGTGTGGACCTCGCCCTGAGAGGTCCAGGGCAGCGTGGCCACA 6792
QY 2277 TATGGCCCATATTTCTGCTGCTTTGAGGTGACAGCAGAGCTGTGGTCTGAGAAGTGG 2336
Db 6791 TATGGCCCATATTTCTGCTGCTTTGAGGTGACAGCAGAGCTGTGGTCTGAGAAGTGG 6732
QY 2337 GGCTCAGGTTCAACAGAGGAGGAGTTCAGAGTCCATATGGCCCAAGATGTGCCCTT 2396
Db 6731 GGCTCAGGTTCAACAGAGGAGGAGTTCAGAGTCCATATGGCCCAAGATGTGCCCTT 6672
QY 2397 CATGAGGACTGGGATATCCCCGCTCAGAAAGAGGACTCCACACAGTCTGGCTGTCC 2456
Db 6671 CATGAGGACTGGGATATCCCCGCTCAGAAAGAGGACTCCACACAGTCTGGCTGTCC 6612
QY 2457 CTTTGTAGTACTCTAGGGGAGGACAGATCAGGATGGCGGTATGTTCCATCTTCACTTGT 2516
Db 6611 CTTTGTAGTACTCTAGGGGAGGACAGATCAGGATGGCGGTATGTTCCATCTTCACTTGT 6552

QY 2517 ACCACAGCAGAAAGTTGGGGGGCCCTCAGGAGATGGGCTTTGGGTAAAGGGGGAT 2576
DB 6551 ACCACAGCAGAAAGTTGGGGGGCCCTCAGGAGATGGGCTTTGGGTAAAGGGAT 6495
QY 2577 GTCCTACTCATGTACAGGAATTTGGGGTTGAGGAGCAGAGCGCTGGCAGGAATAAGAT 2636
DB 6494 GTCCTACTCATGTACAGGAATTTGGGGTTGAGGAGCAGAGCGCTGGCAGGAATAAGAT 6435
QY 2637 GAGTGAGACAGACAGGCTATTGGAATCCACACCCCAAGGAGGCTCAGCCCTGGA 2696
DB 6434 GAGTGAGACAGACAGGCTATTGGAATCCACACCCCAAGGAGGCTCAGCCCTGGA 6375
QY 2697 CACCTCACCCAGGATGTGGCTTCTTTTCACTCCTGTTTCCAGATCTGGGCGAGGTAGG 2756
DB 6374 CACCTCACCCAGGATGTGGCTTCTTTTCACTCCTGTTTCCAGATCTGGGCGAGGTAGG 6316
QY 2757 ACCTCATTTCTCAGAGGTGACTCAGGTCAAGTATAGGACCCCTATCTGGTCTAAAGACAG 2816
DB 6315 ACCTCATTTCTCAGAGGTGACTCAGGTCAAGTATAGGACCCCTATCTGGTCTAAAGACAG 6256
QY 2817 AGCGTCCACAGATCTGCATGCGTTGCGGTGAGGAACATGAGGAGGACTGAGGGTACC 2876
DB 6255 AGCGTCCACAGATTTGCCATGCGTTGCGGTGAGGAACATGAGGAGGACTGAGGGTACC 6196
QY 2877 CCAGGACAGAACACTGAGGAGACTGCACAGAAATCAGCCCTGCCCTGCTGCACCCC 2936
DB 6195 CCAGGACAGAACACTGAGGAGACTGCACAGAAATCAGCCCTGCCCTGCTGCACCCC 6136
QY 2937 AGAGAGATGGGCTGGGCGCTGCGGAGGTCTTCCGTTATCTGGGATCATTTGATGTC 2996
DB 6135 AGAGAGATGGGCTGGGCGCTGCGGAGGTCTTCCGTTATCTGGGATCATTTGATGTC 6076
QY 2997 AGGACGGGAGGCTTGGCTGAGAGGCTGCGCTCAGGTCAAGTATAGGAGGAGGCTCCCA 3056
DB 6075 AGGACGGGAGGCTTGGCTGAGAGGCTGCGCTCAGGTCAAGTATAGGAGGAGGCTCCCA 6016
QY 3057 GCGCCTGCCAGGAGTCAAGGTGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3116
DB 6015 GCGCCTGCCAGGAGTCAAGGTGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5956
QY 3117 CAATGAATTTGATATCTTCTGCTGCGCTTCCCAAGACCTTCCCAAGACCTTCCCAAG 3176
DB 5955 CAATGAATTTGATATCTTCTGCTGCGCTTCCCAAGACCTTCCCAAGACCTTCCCAAG 5896
QY 3177 GTTGTCCCTTCTGCTTCCCTTCTTATCATGATGTTGAATCTTGTATTTGGATTTTC 3236
DB 5895 GTTGTCCCTTCTGCTTCCCTTCTTATCATGATGTTGAATCTTGTATTTGGATTTTC 5836
QY 3237 TCAGACAGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3296
DB 5835 TCAGACAGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5776
QY 3297 AGACAGAGGGGTCATCCACTGATGAGATGGGATGTCACAGATCCAGGAGGAGGAGGAGG 3356
DB 5775 AGACAGAGGGGTCATCCACTGATGAGATGGGATGTCACAGATCCAGGAGGAGGAGGAGG 5716
QY 3357 CTTGTAGCTAGAAAGCAGGAGGCTGCTTGGGTCGACCCCTCAGGAGGAGGAGGAGGAGGAGG 3416
DB 5715 CTTGTAGCTAGAAAGCAGGAGGCTGCTTGGGTCGACCCCTCAGGAGGAGGAGGAGGAGG 5656
QY 3417 TCCTCTTCTGAGCTCCAGGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3476
DB 5655 TCCTCTTCTGAGCTCCAGGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5596
QY 3477 GGTACAGAGAGGAGGATGACAGAGGTTGTCAGAGTGAATTTGTCCTGATATGACACA 3536
DB 5595 GGTACAGAGAGGAGGATGACAGAGGTTGTCAGAGTGAATTTGTCCTGATATGACACA 5536
QY 3537 CCAAGGGCCCACTGTCACAGGACACATAGGACTTCCACAGAGTCTGCGCTCACCTCCCT 3596
DB 5535 CCAAGGGCCCACTGTCACAGGACACATAGGACTTCCACAGAGTCTGCGCTCACCTCCCT 5476

QY 3597 ACTGTAGTCTCTGTAGAAATCGAACCTCTGCTGGCCGGCTGTACCTCTGAGTACCTCTCACT 3656
DB 5475 ACTGTAGTCTCTGTAGAAATCGAACCTCTGCTGGCCGGCTGTACCTCTGAGTACCTCTCACT 5416
QY 3657 TCCTCTTCTAGGTTTTCAGGGAGCAGGCAACCCAGAGGACAGGATTCCTCTGGAGGCCAC 3716
DB 5415 TCCTCTTCTAGGTTTTCAGGGAGCAGGCAACCCAGAGGACAGGATTCCTCTGGAGGCCAC 5356
QY 3717 AGAGGACCAACCAAGAGAAATCTGTAAGTAGGCTTTGTTAGAGTCTTCAAGGTTCACT 3776
DB 5355 AGAGGACCAACCAAGAGAAATCTGTAAGTAGGCTTTGTTAGAGTCTTCAAGGTTCACT 5296
QY 3777 TCTCAGCTGAGGCTCTCACACACTCTCTCTCCAGGCTCTGGTCTCTTATTTGCC 3836
DB 5295 TCTCAGCTGAGGCTCTCACACACTCTCTCTCCAGGCTCTGGTCTCTTATTTGCC 5236
QY 3837 AGCTCTGCCACACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3896
DB 5235 AGCTCTGCCACACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5176
QY 3897 GGAGTCTGCACTGCAAGCCTGAGGAAGCCTTTGAGGCCCAACAGAGGCTCTGGGCTGG 3956
DB 5175 GGAGTCTGCACTGCAAGCCTGAGGAAGCCTTTGAGGCCCAACAGAGGCTCTGGGCTGG 5116
QY 3957 TGTGTGTGCAAGGCTGCAAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4016
DB 5115 TGTGTGTGCAAGGCTGCAAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5056
QY 4017 TGCCACTGCTGGTCAACAGATCTCTCCAGAGTCTCTCAGGAGGCTCTCGGCTTTCCCA 4076
DB 5055 TGCCACTGCTGGTCAACAGATCTCTCCAGAGTCTCTCAGGAGGCTCTCGGCTTTCCCA 4996
QY 4077 CTACCATCACTTCACTCGACAGGCAACCCAGTGGGTTCCAGAGGCTCTGAGGAGG 4136
DB 4995 CTACCATCACTTCACTCGACAGGCAACCCAGTGGGTTCCAGAGGCTCTGAGGAGG 4936
QY 4137 AGGGGCCAAGCACCCTTGTATCTCTGGAGTCTTGTTCGAGCAGTATCACTAAGAGG 4196
DB 4935 AGGGGCCAAGCACCCTTGTATCTCTGGAGTCTTGTTCGAGCAGTATCACTAAGAGG 4876
QY 4197 TGGCTGATTTGGTTGGTTTCTGCTCTCAATATCGAGGAGGAGGAGGAGGAGGAGG 4256
DB 4875 TGGCTGATTTGGTTGGTTTCTGCTCTCAATATCGAGGAGGAGGAGGAGGAGGAGG 4816
QY 4257 CAGAAATGCTGGAGAGTGTATCAAAATTTACAAGACTGTTTCTCTGAGATCTTCGCA 4316
DB 4815 CAGAAATGCTGGAGAGTGTATCAAAATTTACAAGACTGTTTCTCTGAGATCTTCGCA 4756
QY 4317 AAGCCTCTGAGTCTCTGAGCTGGTCTTGGCATTGAGTGAAGGAGGAGGAGGAGGAGG 4376
DB 4755 AAGCCTCTGAGTCTCTGAGCTGGTCTTGGCATTGAGTGAAGGAGGAGGAGGAGGAGG 4696
QY 4377 GCGACTCTATGCTCTTGTACCTGCTAGTCTCTCTATGAGGAGGAGGAGGAGGAGGAGG 4436
DB 4695 GCGACTCTATGCTCTTGTACCTGCTAGTCTCTCTATGAGGAGGAGGAGGAGGAGGAGG 4636
QY 4437 ATCAGATCATGCCCCAAGACAGGCTTCTGATATTTGCTGGTCTATGATTCGAATGGAGG 4496
DB 4635 ATCAGATCATGCCCCAAGACAGGCTTCTGATATTTGCTGGTCTATGATTCGAATGGAGG 4576
QY 4497 GCGGCCATGCTCTCTGAGGAGGAAATCTGGGAGGAGTCTGAGTGTGATGAGGAGTGTATG 4556
DB 4575 GCGGCCATGCTCTCTGAGGAGGAAATCTGGGAGGAGTCTGAGTGTGATGAGGAGTGTATG 4516
QY 4557 GGAGGAGCAGAGTCTCTATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4616
DB 4515 GGAGGAGCAGAGTCTCTATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4456
QY 4617 AAAAGTACCTGGAGTGA-CGGCAGGTGCGGACAGTGTATCCGACGCTATGAGTTCCTGT 4675
DB 4455 AAAAGTACCTGGAGTACCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4396
QY 4676 GGGGTCCAAAGGGCCCTCTGCTGAAACAGGAGTATGTGAAAGTCTTTGAGTATGTGATCAAGG 4735

|||||
Db 4395 GGGGTCCAAAGGGCCCTCGCTGAAACACAGCTATGTGAAAGCCCTTGAGTATGTGATCAAGG 4336
|||||
QY 4736 TCAGTGAAGAGTTCGGTCTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGAGG 4795
|||||
Db 4335 TCAGTGAAGAGTTCGGTCTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGAGG 4276
|||||
QY 4796 AAGAGGAGTCTGAGCATGAGTTGACGCCAAGCCAGTGGAGGGGAGCTGGGCCAGTGC 4855
|||||
Db 4275 AAGAGGAGTCTGAGCATGAGTTGACGCCAAGCCAGTGGAGGGGAGCTGGGCCAGTGC 4216
|||||
QY 4856 ACCTTCAGGGCCGCTCCAGCAGCTTCCCTGCTCGTGTGACATGAGGCCCATCTTC 4915
|||||
Db 4215 ACCTTCAGGGCCGCTCCAGCAGCTTCCCTGCTCGTGTGACATGAGGCCCATCTTC 4156
|||||
QY 4916 ACTCTGAAGAGAGCGGTGAGTCTCAGTAGTAGTCTTCTGTTCTATTTGGGTGACTTGG 4975
|||||
Db 4155 ACTCTGAAGAGAGCGGTGAGTCTCAGTAGTAGTCTTCTGTTCTGTTCTGTTGGGTGACTTGG 4096
|||||
QY 4976 AGATTATCTTTGTTCTCTTTTGAATTTGTTCAAAATGTTTAAAAAGGATGGTTGAA 5035
|||||
Db 4095 AGATTATCTTTGTTCTCTTTTGAATTTGTTCAAAATGTTTAAAAAGGATGGTTGAA 4036
|||||
QY 5036 TGAACITCAGCATCCCAAGTTATGAATGACAGCAGTACACAGCTTCTGTGTATAGTAT 5095
|||||
Db 4035 TGAACITCAGCATCCCAAGTTATGAATGACAGCAGTACACAGCTTCTGTGTATAGTAT 3976
|||||
QY 5096 AAGGTAAGAGTCTGTGTTTATTTCAGATTGGGAATCCATCTATTGTTGAAATGGG 5155
|||||
Db 3975 AAGGTAAGAGTCTGTGTTTATTTCAGATTGGGAATCCATCTATTGTTGAAATGGG 3916
|||||
QY 5156 ATAATACAGCAGTGGAAATAGTACTTTAGAAATGTGAAAATGAGCAGTAAATAGATGA 5215
|||||
Db 3915 ATAATACAGCAGTGGAAATAGTACTTTAGAAATGTGAAAATGAGCAGTAAATAGATGA 3856
|||||
QY 5216 GATAAAGAACTAAAGAAATTAAGAGATAGTCAATCTTGCTGTATACCTCAGCTATTCT 5275
|||||
Db 3855 GATAAAGAACTAAAGAAATTAAGAGATAGTCAATCTTGCTGTATACCTCAGCTATTCT 3796
|||||
QY 5276 GTAAATTTTAAAGATATATGATACCTGATTTCTGCTTCTTGGATGTAAGA 5335
|||||
Db 3795 GTAAATTTTAAAGATATATGATACCTGATTTCTGCTTCTTGGATGTAAGA 3736
|||||
QY 5336 GAAATTAATCTGAATAAAGAAATTTCTCTGTTCACTGGCTTTCTTCTCCATGCACT 5395
|||||
Db 3735 GAAATTAATCTGAATAAAGAAATTTCTCTGTTCACTGGCTTTCTTCTCCATGCACT 3676
|||||
QY 5396 GAGCATCTGCTTTTGAAGGCCCTGGGTTAGTAGTGAGATGCTAAGGTAAGCCAGCT 5455
|||||
Db 3675 GAGCATCTGCTTTTGAAGGCCCTGGGTTAGTAGTGAGATGCTAAGGTAAGCCAGCT 3616
|||||
QY 5456 CATACCCACCATAGGCTGCTAGAGTCTAGAGCTGAGTCACTGATATCAGGTTGGCAAG 5515
|||||
Db 3615 CATACCCACCATAGGCTGCTAGAGTCTAGAGCTGAGTCACTGATATCAGGTTGGCAAG 3556
|||||
QY 5516 ATGTCTCTTAAGAGTGTAGGAAAAGTGTAGAGAGGGGTGAGGGGTGCGGGCTCCGGGTGA 5575
|||||
Db 3555 ATGTCTCTTAAGAGTGTAGGAAAAGTGTAGAGAGGGGTGAGGGGTGCGGGCTCCGGGTGA 3496
|||||
QY 5576 GAGTGTGGAGTGTCAATGCCCTGAGTGGGGCAATTTGGGCTTTGGGAACTGCAGTTC 5635
|||||
Db 3495 GAGTGTGGAGTGTCAATGCCCTGAGTGGGGCAATTTGGGCTTTGGGAACTGCAGTTC 3436
|||||
QY 5636 CTCTGGGGAGCTGATGTGTAATGATCTTGGGTGGATCC 5674
|||||
Db 3435 CTCTGGGGAGCTGATGTGTAATGATCTTGGGTGGATCC 3397

RESULT 4

AF134576 AF134576 49375 bp DNA PRI 06-APR-1999
LOCUS Homo sapiens MAGE10 and undefined MAGE genes, complete sequence.
DEFINITION AF134576
ACCESSION

94567140
AF134576.1 GI:4567140
HTG.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
Sequence of three cosmid of human Xq28 (16c3, 3g5, 3h5) containing
MAGE genes
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 49375)
AUTHORS
Gloeckner, G., Rosenthal, A. and Schattevoy, R.
TITLE
Direct Submission
JOURNAL
Submitted (12-MAR-1999) Genome Analysis, Institute for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
Location/Qualifiers
1..49375
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="Xq28"
/chromosome="X"
/clone="cosmids 16c3; 3g5; 3h5"
1..10000
misc_feature
/note="incomplete at 3' end"
repeat_region
/note="Region: Cosmid 16c3"
/rpt_family="MER93"
repeat_region
/evidence-not_experimental
289..372
/rpt_family="LIM1"
repeat_region
/evidence-not_experimental
446..726
/rpt_family="LIM1"
repeat_region
/evidence-not_experimental
528..43599
misc_feature
/note="Region: Cosmid 3g5"
repeat_region
/rpt_family="LIMC3"
731..4392
repeat_region
/evidence-not_experimental
complement(4393..4696)
/rpt_family="MER2"
repeat_region
/evidence-not_experimental
complement(4697..4755)
/rpt_family="LIM2"
repeat_region
/evidence-not_experimental
complement(4756..4803)
/rpt_family="MER2"
repeat_region
/evidence-not_experimental
4805..4847
/rpt_family="LIMC3"
repeat_region
/evidence-not_experimental
4837..5414
/rpt_family="LIMC5"
repeat_region
/evidence-not_experimental
5566..6545
/rpt_family="LIM4"
repeat_region
/evidence-not_experimental
complement(6505..6963)
/rpt_family="MIR"
repeat_region
/evidence-not_experimental
7671..7697
/note="111 with 85% homology to 112"
repeat_region
/rpt_type-inverted
/evidence-not_experimental
7814..7981
misc_feature
/note="homology = 66.10%, score = 48, counts = 28"
/rpt_type-tandem
/rpt_unit-cctcac
/evidence-not_experimental
7814..8708
/note="GC score = 44.00 (895bp)"

QY 2377 GGCCCAAGATGTGCCCCCTTCTCAGAGACTGGGGATATCCCCGCTCAGAAAGAGGAC 2436
DB 29585 GCCCAATGTGTGCCCCATTATGAAGATGGGGATA-CCTTGGCTCAGAAAGAGGAC 29643
QY 2437 TCCACAGACTGTGCTGCTCCCTTTTAGTAGCTCTAGGGGACCACGATCAGGATGGCG 2496
DB 29644 CCCACAGACTGTGCTGCTCCCTGATTTTGTCTCAGAGGGGACCAATCAGGATAGCCC 29703
QY 2497 TATGTTCCATTCTCATTGTACACAGGAGAAATTTGGGGGCCCTCAGGAGATGGG 2556
DB 29704 TATGTGCCAACCTCATTTGTGCCACAGAAAGATTTAGAGAGCCCTCAGGATGGG 29763
QY 2557 TCTTGGGTAAGGGGGATGTCTACTCATGTCTCAGGAAATTTGGGGTTGAGGAACACAG 2616
DB 29764 TCTTGCAGTAAGGGGAGCTATCTGCTCATCTCAGGGGTTTTCAGGTTGAGGATGCG 29823
QY 2617 GCGTGGCAGGATAAAGATGAGTGAGACAGACAAAGGCTATTGGAATCCACACCCAGAA 2676
DB 29824 GCGCCATCAGC-ATGAAGAGTAAACCCACAGAGGAGCCATAGAAACACTCACCCAGAA 29879
QY 2677 CCAAGGGGTGAGCCCTGGACACCTCA-CCAGGATGTGGCTTCTTTTTCAC 2727
DB 29880 CCAAGGGGTATACCTTGGACACCCCATGTGGGGGTACAGGATGTAGC-TCCATCTCAT 29938
QY 2728 TCTGTTTCCAGATCTGGGGCAGGTGAGGACCTCATTTCTCAGAGGGTGACTCAGGTCAAC 2787
DB 29939 TCTGTTTTCAGATCTCGGGAGGTGAGGAACTTGTTCTCCGAGGATGACTCAGGTCAAC 29998
QY 2788 GTAGGACCCCATCTGGTCTAAGACAGAGCGGTCCAGGATCTGCCATGCTTGGGT 2847
DB 29999 ACAGGGCCCCCATCTGGTGATAGACAGAGTGGTCCAGGATCTGTCACTAGTTCCGGT 30058
QY 2848 GAGGAACATGAGGAGGACTGAGGATACCCAGGACAGAACACTGAGGAGACTGCACA 2907
DB 30059 GAGGAACATGAGGAGGACTGAGGAGACCTTTGGGCCAGAACAGATGAGGACCTCAG 30118
QY 2908 GAAATAGCCCTGCCCTGTGTACCCACAGAGACATGGGCTGGGCCGTCTGCCAGGT 2967
DB 30119 GAAATCTGCCCTGCCCTGTGTACCTCCAGAGAGCATGGGAGGCTGTCTGCTGAGT 30178
QY 2968 CTTTCC-GTTATCTGGGATCATTTGATGATCAGGAGCGGGAGGCTTGTCTGAGAGG 3025
DB 30179 CCCCCCATTACCTGGGATATGTTGTGTCAGGAGTGGGAGGCTTTGTC-GAGGGGT 30237
QY 3026 CTGCGCTCAGGTGAGTGGAGGCTGCCAGGCCCTGCCAGGAGTCAAGGTGAGGACCA 3085
DB 30238 CTGCACTCAGGTGAGTGGAGGCTTTAGGCCCTGCCAGGACAGAGTGAAGACGA 30297
QY 3086 AGGGGCACTCACCAGGACATTAATTTCCAAATGAATTTTGAATFCTCTTCTGCTCCCT 3145
DB 30298 AGCAGGTTCTCACCAGGACATGAATTTCCAAATGATTTTCAAGTCTCTTCTGCTCT 30357
QY 3146 TCCCCA-AGGACCTAGGACGTTGGCCAGATGTTTGTCCCTCTCTCTCTCTCTCTCT 3204
DB 30358 TCCCAAGAGGACCTGGGACGTTGGCCAGATGTGAGTCTCTCTCTCTCTCTCTCTCT 30414
QY 3205 TATCATGATGTGAACCTTTGATTTGATTTCTCAGACCAAGAAAGGGCAGGATCCAGG 3264
DB 30415 TATCAGGATGTGAGCTTTAATCTGAGTTCTCAGGCCAGAAAGGGTGGATCCAGG 30474
QY 3265 CCTTCCAGGAAATAAAGGGCCCTGCTGAGACAGAGGGGGTATCATCCACTGCATGA 3324
DB 30475 CCTTCCAGGAAAGGTGAGGGCCCTGTGTGAGCAGAGGGGACCATTCACCCCAAGA 30534
QY 3325 GAGTGGGATGTACAGAGTCCAGGCCACCTCTCTGTAGACACTGAGAACCCAGGCTGT 3384
DB 30535 GGTGGGAGACCTCACAGATTTCCAGCTTACCTCTCTGTGTAGACTGGGGGCTCTGAGGCT 30594
QY 3385 GCTTGGGCTGTGACCCCTGAGGGCCCTGGATTTCTCTCTCTCTCTCTCTCTCTCTCT 3444
DB 30595 GCTTGCAGTCTGCACCTGAGGGCCCATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 30654

QY 3445 CAGTGGGCTTGGTCTGAGACAGTATCTCTCAGGTACAGAGCAGAGGATGTCACAG- 3501
DB 30655 CACTGAGGCTTGGTCTGAGCGCGTCCCTCAGTTCACAGAGCAGAGGAGATGCAGAGCT 30714
QY 3501 GGTGTGCCAGCAGTGTATGTTTGGCCCTGAATGACACCAAGAGGGCCCACTCCACAGGA 3560
DB 30715 CTAGTCCAGCAGTGAACGTTTGGCTTGAATGACACTAATAGCCCCCATCGCCCCAGAA 30774
QY 3561 CACATAGGACTCCACAGAGTCTGGCTCA-CCTCCCTACTGTCTCAGTCTCTAGAAATGAC 3619
DB 30775 CATATGGGACTCCAGAGCAGTCTGGCTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 30834
QY 3620 CTCTGTGGCGGCTGTACCTCTGACCTCTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3679
DB 30835 CTCTGTGTCTGTATACCTGAGGTGCTCTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 30894
QY 3680 CAGGCCAACCCACAGAC-AGGATTCCTGGAGGCCACAGAGGAGCAC 3726
DB 30895 CAGGCTGACCAAGATCACCAGGAAGCTCCAGAGGATCCCCAGGAGGCCCTAGAGGAGCAC 30954
QY 3727 C-AAGGAGAAGTCTGTAACTAGGCTTTTGTAGAGTCTCCTAAGGTTTCTCAGTCTCAGCTG 3785
DB 30955 CAAGGAGAAGTCTGTAACTAGGCTTTTGTAGAGCTCTCCTAAGGTTTCTCAGTCTCAGCTG 31014
QY 3786 AGGCTCTCACACACTCCCTCTCTCCCGAGGCTGTGGGTCTTCTATTTGCCAGTCTCTCTG 3845
DB 31015 AGGCTCTCACACTCTCTCTCTCTCCAGGCGAGTGGGTCTCTCTCTCTCTCTCTCTCTCTCT 31074
QY 3846 CCACACT 3905
DB 31075 CCACACT 31134
QY 3906 ACTGCAAGCTGAGGAGGCTTTGAGGCCCAAGAGGCGCTGGGCTGGTGTGTGTG 3965
DB 31135 ACTGCAAGCTGAGGAGGCTTTGACACCAAGAGGCGCTGGGCTGGTGTGTGTG 31194
QY 3966 AGGCTGCCAC-TCCT 4004
DB 31195 AGGCTGCCACTACTGAGGAGCAGGAGGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 31254
QY 4005 CCCTGAGGAGGCTGCCACTGTGGGTCAACAGATCTCTCCAGAGTCTCTCAGGAGGCTCT 4064
DB 31255 CCCTGAGGAGGCTGCCACTGTGGGTCAACAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 31314
QY 4065 CCCTCTTCCACTTACCTTCACTCTCAGAGGAGGCAACCCAGTGAAGGCTTCAGCA 4124
DB 31315 CCCTCTTCCACTTCCACTTCACTCTATGAGGAGGCAACCCAGTGAAGGCTTCAGCA 31374
QY 4125 GCGTGAAGAGGAGGCGCCCAAGCCTCTTGTATCTCTGGAGTCTCTTCTCTCTCTCTCTCTCTCTCT 4184
DB 31375 ACCAAGAAGAGGAGGCGCCCAAGCCT 31434
QY 4185 TCACTAAGAGGCTGCTGATTTGGTGTCT 4244
DB 31435 TCACTAAGAGGCTGCTGATTTGATTTCT 31494
QY 4245 CAGTCAAAAGGAGGAGGCTGCTGAGAGTCTCATCAAAATTAACAGCACTGTTTCTCTCTCTCTCT 4304
DB 31495 CGGTCAAAAGGAGGAGGCTGCTGAGAGGCTCATCAAAATTAACAGGCTGCTTCTCTCTCTCTCTCT 31554
QY 4305 AGATCTTCGGCAAGGCTCTGAGTCT 4364
DB 31555 AGATCTTCGGCAAGGCTCTGAGTCT 31614
QY 4365 CAGACCCACCGGCT 4424
DB 31615 CAGACCCACCGGCT 31672
QY 4425 TGTCTGGTGTATATCAGATCATGCCCAAGAGGCTTCTCTGATAATTTCTCTCTCTCTCTCTCTCTCTCT 4484
DB 31673 TGTCTGGT--TATCAGATCATGCCCAAGAGGCTTCTCTGATAATTTCTCTCTCTCTCTCTCTCTCTCT 31730
QY 4485 TTGCAATGAGGCGGCCCATGCT 4544

|||||
Db 31731 TTGCAATGGAGGCAAAATCGCGTCCCTGAGGAGAAATCTGGGAGGAGCTGAGTGTGATGA 31790
QY 4545 AGGTGTATGATGGGAGGAGCAGACAGTGCCTATGAGGAGGCCAGGAGCTGCTCACCCCAAG 4604
Db 31791 AGGTGTATGAGGAGGAGCAGACAGTGTCTGGGAGGCCAGGAGCTGCTCACCCCAAG 31850
QY 4605 ATTTGGTCAGAGAAAGTACCTGGAGTA-CGGCAGGTGCCGAGCAGTGAATCCGCACGCT 4663
Db 31851 ATTTGGTCAGAGAAATACCTGGAGTACCGGAGGAGTGCAGCAGTGAATCCCATATGCT 31910
QY 4664 ATGAGTCTCTGGGGTCCAGGCGCTCGCTGAACACAGCTATGTGAAGTCTTGTGAGT 4723
Db 31911 ATGAGTCTCTGGGGTCCAGGCGCTCGCTG-----CTTGAAGTACTGGAGC 31960
QY 4724 ATGTGATCAAGTCAAGAGTTCGGTCTTTTCTTCCATCCCTGCGTGAAGCAGCTT 4783
Db 31961 AGGTGTAGGTCAATGCAAGAGTCTCATTTCTTACCCTCCCTGCGTGAAGCAGCTT 32020
QY 4784 TGAGAGAGGAGGAGGAGTCTGAGCATGAGTTGCGAGCCAGGCGAGTGGAGGGGA 4843
Db 32021 TGAGAGAGGAGGAGGAGTCTGAGCATGAGTGCAGGAGGCGCAGTGGAGGGGG 32080
QY 4844 CTGGCCAGTGCACCTTCAGGCGCGCTCCAGCAGCTTCCCTGCC-TCGTTGACATG 4902
Db 32081 CTGGCCAGTGCACCTTCAGGCGCGCTCCAGTATTTCCCTGCTTAAATGTGACATG 32140
QY 4903 AGGCCCATTTCTCACTC--TGAAGAGAGGGTCAAGTCTTCTCAGTAGTAGTTCGTGTC 4960
Db 32141 AGGCCCATTTCTCTCTTTGAAGAGAGCAGTCAACATTTCTAGTAGTGGTTCGTGTC 32200
QY 4961 TATTGGGTGACTGGAGATTTATCTTTGTTCTTTTGGAAATGTTCAAAATGTTTTTTT 5020
Db 32201 TATTGGGTGACTTTGAGATTTGCTTTGTTCTTTTGGAAATGTTCAAAATGTTTCT 32259
QY 5021 TAAGGATGGTTGAATGAATTCAGCATCCAGTCTTATGAATGAGCAGTGCACAC--AG 5078
Db 32260 TAAGGATGGTTGAATGAATTCAGCATTCAAATTTATGAATGAGCAGTGCACACATG 32319
QY 5079 TTCTGTATATAGTTTAAAGGTAAGTCTTGTGTTTATTCAGATTTGGGAAATCCCAT 5138
Db 32320 TGCTGTATATAGTTTAGGAGTAAAGTCTGTTTATTCAGATTTGGGAAATCCCAT 32379
QY 5139 CTATTTTGTGAATTTGG--ATAATTAACAGCAGTGAATTAAGTACTTA-GAAATGTGA 5195
Db 32380 CCAATTTGTGAATTTGGACATAGTTACAGCAGTGAATTAAGTACTTAAGAAATGTA 32439
QY 5196 ATGACAGTAAATAGATGAGATAAGAACTAAAGAAATTAAGAGATAGTCAATTTCTGC 5255
Db 32440 ATGACAGTAAATAGTATGACA-----TAAAGAAATTAAGAAATTTAATTTCTGC 32491
QY 5256 CTATACCTCAGCTATTCTGTAAA--TTTTTAAAGATATATGATACCTGGATTTCTCT 5313
Db 32492 CTATATA-CTCAGTCTATTTCGTTAAATTTTAAAGATTTGATACCTGGATTTCTCT 32550
QY 5314 TGGCTTTCTTGAATTTGAAGAAATTTAAATCTGAATTAAGAAATTTCTTCTGTTCACTG 5373
Db 32551 TGGCTTTCTTGAATTTGAAGAAATTTAAATCTGAATTAATCAATTTCTCCTGTTCACTG 32610
QY 5374 GCTCTTTCTTCTCCTCAGCTAGCATCTGCTTTTGAAGGCGCTGGGTAGTAGTGG 5433
Db 32611 GCTCAATTTATTTCTATGCACTGAGCATTTGCTGTGGAAGGCGCTGGGTAAATAGTGG 32670
QY 5434 AGATCTTAAGTTAAGCAGACTCATACCCAGTGGTCTGATAGTCTAGGAGCTGCA 5493
Db 32671 AGATCTTAAGTTAAGCAGACTCATACCCAGTGGTCTGATAGTCTAGGAGCTGCA 32730
QY 5494 GTCAGTAAATCGAGTGGCAAGATCTCTCTAAAGATGTAGGAAAGTGAAGAGGGGT 5553
Db 32731 GTCATATAATTAAGTGGAGAGATGCCCTCTTAAGATGTAGAGAAAGTGAAGAGGGT 32790
QY 5554 GAGGTGTGGGCTCCGGGTGAGTGGTGGAGTGTCAATGCCCTCAGCTGGGGCATTTT 5613
|||||

Db 32791..GAGGCTGTGGGCTCCAGGTGAGAGTGGTCCAGAGGTAATCCCTGTGTGGGCGCTTTT 32850
QY 5614 GGGCTTTGGGAACTGCAGTCTCTTCTGGGAGGAGTGTAAATCATCTTTGGGTGG 5670
Db 32851 GGACITTTGGGAACTGCAGTCTTCTCGGAGGAGCTGATCTTAATGAATCGCGGGG 32907
RESULT 5
LOCUS HSU10687 11495 bp DNA PRI 23-JUN-1995
DEFINITION Human MAGE-4a antigen (MAGE4a) gene, complete cds.
ACCESSION U10687
NID 9533514
VERSION U10687.1 GI:533514
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 11495)
AUTHORS De Plaen,E., Arden,K., Traversari,C., Gaforio,J.J., Szikora,J.P.,
De Smet,C., Brasseur,F., van der Bruggen,P., Lethé,B., Lurquin,C.
Brasseur,R., Chomez,P., De Backer,O., Cavenée,W. and Boon,T.
Structure, chromosomal localization, and expression of 12 genes of
the MAGE family
Immunogenetics 40 (5), 360-369 (1994)
JOURNAL MEDLINE 95012457
REFERENCE 2 (bases 1 to 11495)
AUTHORS De Plaen,E.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-1994) Etienne De Plaen, Ludwig Institute for
Cancer Research, 74 Avenue Hippocrate, Brussels, 1200, Belgium
FEATURES
source location/Qualifiers
1..11495
/organism="Homo sapiens"
/isolate="patient M2"
/db_xref="taxon:9606"
/chromosome="X"
/sex="female"
/cell_type="lymphocyte"
/tissue_type="blood"
/dev_stage="adult"
450..510
/note="exon 1-1"
836..937
/note="exon 1-2"
1418..1556
/note="exon 1-3"
2108..2246
/note="exon 1-4"
2836..2928
/note="exon 1-5"
3508..3607
/note="exon 1-6"
4906..4963
/note="exon 1-7"
5519..5648
/note="exon 1-8"
9377..9448
/number=2
9524..11091
/number=3
9589..10542
/gene="MAGE4a"
9589..10542
/gene="MAGE4a"
/codon_start=1
/product="MAGE-4a antigen"
/protein_id="AAAG8871.1"
/db_xref="PID:g533515"
/db_xref="GI:533515"
/translation="MSSEKQSKHCKPEGVQAQEAALGLVGAQPTTEQEAAYSSSS
PLVPGLTEVPAAESAGPPQSPQASALPTTISTCWRQPNEGSSSQEEGPTSPDA

ESLFRALSNNKDELAFHLRKYRAKELVTKAEMLSERVKNYKRCFPVIFGKASESLK
MIFGIDVKEYPASNTYTLVTCLGSLYDGLGNQIIPFKTGLLIIVLGTIAMGDSAS
EEIWEELGVYGVDRHTYVGEPRKLLTQDWQENLEYRQVPGNSPARYEFLWGP
RALAETSIVKVLHVVRNVRVRIAPSLRREALLEEVEG"
polya_signal 11072..11077
polya_site 11091
BASE COUNT 2530 a 3788 c 2924 g 2253 t
ORIGIN

Query Match 46.88; Score 2655; DB 9; Length 11495;
Best Local Similarity 73.68; Pred. No. 0; Mismatches 1100; Indels 416; Gaps 50;
Matches 4231; Conservative 0

Qy 21 CCTCCCTTACACCCCAATCCCTCCCTTTACAGCCACCCATCCAAACATCTTCACGCTC 80
Db 5969 CCAGCACCCCTATCTCCCAACACCCCACTACCTTATGCTTCATCCCAACCCCAAC 6028

Qy 81 ACCCCAGCCCAAGCAGGAGAGATCCGGTTCCACCCTGCTCTCAACCCAGGAAGC-C 139
Db 6029 ACCACTATCCCATCCAGGTTGAATCGCATTCCTGCTTTCAACCCAGGAAGCTC 6088

Qy 140 CAGGTGCCAGATGAGCCGACATGACTTGAAGCATTAGTGTAGAGAGAGCGAGTTT 199
Db 6089 CAGGTTCTGTGATGACGAGTGTGCAATGGGGTTAGAGAGAGCTAGGTTT 6148

Qy 200 TCGGTCTGAGGGGGGCTTGAGATCGGTGGAGGAAGCGGCCCA-GCTCTGTAAAGAGG 258
Db 6149 TCAGTCTGACAGGAGCTTGGATTGGCAGAGGAAGCGGTCCAGGCTCTGTAGGTTG 6208

Qy 259 CAAGGTGACATGCTGAGGAGGAGT- - -GAGGACCCATTTACCCAGATAGAGAGCCCA 315
Db 6209 CATAGTGAAGACTGAGGAGGAAGTGGGAGGCTCTCCACCCAGATAGAGAGCCCA 6268

Qy 316 AATAATCC- - -CTTCATGCCAGTCTGACCATCTGGTGGTGGACTTCTCAGGC 366
Db 6269 AATAATCCGGACCCCTCTCTCCAGTCTGGGCCACCCGTTGGGGGAGCTTCTGAGTC 6328

Qy 367 TGGGCCACCCCGCCCTTGCTGTCTAAACCACTGGGGACTCGAAGTCAAGACTCCGT 426
Db 6329 TGGGACCCCAACCCCACTGCGCTGAAGCCGAGGACTATGGAGTCAAGACTTGGT 6388

Qy 427 GTGATCAGGAAGGCTGCTTAGAGAGGAGCGCTCAGGCTCTGCCAGACATCATGCT 486
Db 6389 GTGATCAGTGCAGGACTGTTGGGGT- - -AGGCTCTGCCAGGCATCAAGT 6436

Qy 487 CAGGATCTCAAGAGGCTGAGGCTCCCTAAGACCCCACTCCCGTGACCAACCCCAAC 546
Db 6437 CAGGACCTTAGGAGGCTGAGTGTCCCCCA- - -CCCCCATTCCTATCCCTTACCCCTT 6493

Qy 547 TCCAATGCTCACTCCCGTGACCAACCCCTCTTTCATTTGTCATTTCAACCCCAACCCCAAC 606
Db 6494 TCCCATCTGCACTCC- - - - -TACCCCATCTGTACCCCC 6527

Qy 607 ATCCCCCAACCCATCCCTCAACCCGTGATGCCCATCCGCCAGCCATTCACCCCTCACCC 666
Db 6528 ATTCACCCACCTGTGCCCTATCTCCCAACCCCAACCCCAACCCCAACCCCAACCCCAAC 6587

Qy 667 CACCCCAACCCCAACCCCACTCCCAACCCCAACCCCAACCCCAACCCCAACCCCAAC 725
Db 6588 CACCCCTACCTTCACTCCCACTAGTGCAGCATCCGGTTCCACCCCTGCTTCAATCCAGG 6647

Qy 726 AAACATCGGGTGCAGGATGTGACGCACTGACTTGGCATTTGGGGCAGAGAGAGC 785
Db 6648 CAACCCCTGGTGGCCGATGTGATGCCACTGACTTGTGAATGAGGGTTAGAGAGT 6707

Qy 786 GAGGTTTCCATTTAGAGGAGCGGTAGAGTTGGCCGAAGAGAACCTGACCCAGGCTCTG 845
Db 6708 GAGTTTCTGGGCTGAAGGGTGGC-TTGAGATCGGACAGAGGAAGGTGGCCAGGCTTG 6766

Qy 846 TGAGGAGCAGGTGAGAGGCTGAGGAGGAGTCTAGGACCCCGCCACTTCCAAATAGAG 905
Db 6767 TGAAGAGGCAAGTGAGACTCTGAGGAGGAGTTCAGGAACCCCTTATCCCTGTATAGAGG 6826

Qy 906 CCCCATAATTTCCAGCCCCCCTTGTGCTGCCAGCCTTGCCACCCCGGGAGACGCTCT 965
Db 6827 TCCAGAGCCTGGACTACC- - - - -CGCGAGGCTGACTTCT 6862

Qy 966 CAGCTGGCTGCCCCCAGACCCCTGCTCCTCAAAAGCCTTGAGAGACACAGGTTCTTCTC 1025
Db 6863 CAGACTGGCTGCTCCCGCAGCTCCGCC- - - - -TTGCAACGCTTGTGTTAAGCCAC 6917

Qy 1026 CCAAGCTCTGGATCAGAGTTGCTGTGACAGGCGAGGACTGTTAGGAGGCGGAGG 1085
Db 6918 AGGGACTCTGGAGTCAGAGTTGGTGTGATCAGGAAGGCTGTTAGGAGA-GGCATG 6976

Qy 1086 GCAGAGCTCTGCCAGGATCAAGATCAGACCCCAAGAGGAGGCTGTGGGCCCAAG 1145
Db 6977 GCCAGGCTGCCAGGAATCAAGTCAGAACCC-TGAGAGGGAATGAGGTCCCGCAAG 7035

Qy 1146 ACTGCACTCAATCCCACTCCAGCCCATTCGATTTCCCATTCCTCCCAACCCCAACCCCAT 1205
Db 7036 ATCTAGTCTAAACCCCACTCCCAAC- - - - - 7063

Qy 1206 CTCTCAGCTACACCTCCACCCCATCCCTACTCTCTACTCCGTACCTGACCCACCT 1265
Db 7063 - - - - -ATCCGCTGCCATTCGCTGCTCCATTTCCCATTCCTCTTCCCT 7104

Qy 1266 CCAGCCCCAGCACCGCCCAACCCCTTCTGCCACCTCACCTCACTGCCCAACCCCAAC 1325
Db 7105 CCACCTTCACT- - - - - 7117

Qy 1326 CTTCTATCTCTCATGTGCCCACTCCCATCGCTTCCCTCCCATTTCTGGCAGAAATCCGGTTT 1385
Db 7117 - - - - -GGCAGAAATCCAGTTT 7131

Qy 1386 GCC-CTGTCTCAACCCAGGAGCCTGTAGGCCCGATGTGAACCACTGACTTGA 1444
Db 7132 CTTCTGCTATCAATCCAGGAACCCAGGCTGTGTGCTGGATTTTT- - - - - 7184

Qy 1445 CCTCAGATCTCAGAGAAGCCAGTTCAATTAATGTTCTGAGGGCGGCTTGAGATCC 1504
Db 7184 - - - - -TGGGGTTCAGAGAAATCAAGGCATCTCTGAGGGCCAGTTGAGATCG 7232

Qy 1505 ACTGAGGGAGTGTGTTAGGCTCTGTGAGGAGCAAGGTGAGATCTGAGGAGGAGTCT 1564
Db 7233 GCTGAGGGAGCGGGCCCAAGCTCTGTGGGAGGCAAGGTGAGACTCTGAGGAAGGACTG 7292

Qy 1565 AGAGGCACACACCCAGGTAGTGGCCCAAAATGATCCAGTACCACCTCTGCTGCCAG 1624
Db 7293 AGGAGCCCCCAACCAAGATAGA-GGAACCAAAATATCCAGCGAGCTCTCTGCTGCCAG 7351

Qy 1625 CCTGTGACCCCGCCAGGACAGATGCTCAGCTGACCAACCCCGCTCCCGTCCCACT 1684
Db 7352 TCTGTGACCCCGG- - -GGGAGACTTCTCA- - -GGCTAGGCCATCCAGCTCCCACT 7404

Qy 1685 GCACTTAACCCACAGGCAATCTGTAGTATAGCTTATGTGACCCGGGAGGTTGGTCT 1744
Db 7405 GCACTTAAGCTACAGGGGACTGTAGAGTCAAGAGTGTGGTGTGCCCA- - - - - 7453

Qy 1745 AGGAGAGGAGGCGGCAATCAAGGTCAGCATCCCGCCGCAATTAGGTCAGGACCC 1804
Db 7453 - - - - -AGGAGGCGGCGG- - - - -TCTGCTGGCATCGGGGCTCAGGACT 7493

Qy 1805 TGGAGGGAATGAGGTTCCCAACCCCACTGCTCTCTCTCATCTCCAGCCCAACCCCAAC 1864
Db 7494 TGAGAGGGAATGAGGCGCTACACCCCAACCCCACTCCGATTC- - - - -CAACAT 7543

Qy 1865 TCACATTTCCATACCTACCCCTACCCCTACCCCAACCTCATCTTGTGAGAAATCCTGCTGTA 1924
Db 7544 GCGAGGCCCATCCCAACCTCCGTTTTCAGAAATCCATTTT- - - - -TCCCTGCACTCA 7599

Qy 1925 CCGAGGAAGCCAGGGAATGGCGGCGAGGCACTCGGATCTTGCAGTCTTCCATCCAGGCT 1984
Db 7600 CCGGGAAGACCTTGGGAATGGT- - - - -CAGGCACTCGGATCTTGACATCCACATCGAGGGC 7656


```
/evidence-not_experimental
complement(1683. .1767)
/note="M2EF, score = 60%"
/evidence-not_experimental
complement(2062. 2120)
/note="FEXHB, score = 11.76%"
/evidence-not_experimental
complement(2637. .2710)
/rpt_family="L1ME2"
/evidence-not_experimental
complement(2859. .3028)
/rpt_family="L1ME3A"
/evidence-not_experimental
complement(2911. .2978)
/note="M2EF, score = 79.7%"
/evidence-not_experimental
4250. 4431
/note="GRAIL, score = 52.000%, comment = good shadow"
/evidence-not_experimental
complement(4639. .4866)
/rpt_family="L1MA2"
/evidence-not_experimental
complement(4893. 5203)
/rpt_family="L1MA2"
/evidence-not_experimental
complement(4974. .5117)
/note="GRAIL, score = 61.000%, comment = good"
/evidence-not_experimental
complement(5206. .5561)
/rpt_family="L1MA2"
/evidence-not_experimental
complement(5562. .5602)
/rpt_family="L1MA9"
/evidence-not_experimental
complement(5619. 5655)
/rpt_family="L1MA5"
/evidence-not_experimental
complement(6084. .6254)
/rpt_family="MLTID"
/evidence-not_experimental
complement(6262. 6322)
/rpt_family="MLTID"
/evidence-not_experimental
complement(6384. .6549)
/rpt_family="MLTID"
/evidence-not_experimental
6664. 6771
/note="homology = 89.80%, score = 82, counts = 27"
/rpt_type=tandem
/rpt_unit=ccctt
/evidence-not_experimental
6807. 6884
/note="FEXHB, score = 10.83%"
/evidence-not_experimental
7337. 7454
/rpt_family="L1"
/evidence-not_experimental
7483. 7533
/note="FEXHB, score = 12.95%"
/evidence-not_experimental
7950. 8010
/rpt_family="MSTA"
/evidence-not_experimental
8147. 8238
/rpt_family="MSTA"
/evidence-not_experimental
8245. 8286
/rpt_family="MSTA"
/evidence-not_experimental
8287. 8361
/rpt_family="MSTC"
/evidence-not_experimental
8878. 8921
```

```
/note="homology = 97.70%, score = 20, counts = 2"
/rpt_type=tandem
/rpt_unit=gattttacctgtcaagtgcta
/evidence-not_experimental
complement(9471. .9526)
/note="M2EF, score = 85.6%"
/evidence-not_experimental
complement(9474. .9529)
/note="FEXHB, score = 12.02%"
/evidence-not_experimental
9618. 9660
/rpt_family="MER3"
/evidence-not_experimental
10066. 10099
/note="GRAIL, score = 61.000%, comment = good"
/evidence-not_experimental
11110. 11190
/note="GRAIL, score = 76.000%, comment = excellent"
/evidence-not_experimental
11257. 11275
/note="FEXHB, score = 19.61%"
/evidence-not_experimental
11360. 11419
/note="homology = 91.70%, score = 20, counts = 2"
/rpt_type=tandem
/rpt_unit=agggcaaatagtagaactatcatcaggc
/evidence-not_experimental
11613. 11637
/note="xpond exon prediction, score = 68% (0%)"
/evidence-not_experimental
complement(12630. .12662)
/note="FEXHB, score = 13.07%"
/evidence-not_experimental
13530. 13600
/rpt_family="MLT2B2"
/evidence-not_experimental
13607. 13760
/rpt_family="MLT2B2"
/evidence-not_experimental
13820. 13947
/note="GRAIL, score = 88.000%, comment = excellent"
/evidence-not_experimental
13924. 13977
/rpt_family="MLT2D"
/evidence-not_experimental
13978. 14043
/note="homology = 75.80%, score = 23, counts = 6"
/rpt_type=tandem
/rpt_unit=ctctctctctt
/evidence-not_experimental
14054. 14087
/note="homology = 94.10%, score = 28, counts = 17"
/rpt_type=tandem
/rpt_unit=tg
/evidence-not_experimental
complement(15048. .15137)
/note="FEXHB, score = 12.14%"
/evidence-not_experimental
complement(15049. .15138)
/note="M2EF, score = 83.8%"
/evidence-not_experimental
15192. 57359
/note="Region: Cosmid IG0799"
complement(15720. .15884)
/note="GRAIL, score = 53.000%, comment = good"
/evidence-not_experimental
16505. 16631
/rpt_family="L1MA9"
/evidence-not_experimental
16659. 16746
/rpt_family="MER4A"
/evidence-not_experimental
16798. 16851
```


Qy	1864	CTCACAATCCCATACCTACCCCTACCCCCAACCTCATCTGTGCAGAAATCCCTGCTGTCA	1923
Db	59076	TGCCAGGCCCATCCCCAACTCCGTTTGCAGAAATCCATTTT-----TCCCTTCGATCA	59131
Qy	1924	ACCCACGGAACCCACAGGAAATGGCGCCAGGCACTCGGATCTTGACCTCCCCATCCAGGG	1983
Db	59132	ACCCCGGAAGACCTGGGAATGGT---CAGGCATCTCGATCTTGACATCCACATCGAGGG	59188
Qy	1984	TCTGATGGAGGAAGG-----GCTTGAACAGGGCTCAGGGAGCAGAGGGAG---	2033
Db	59189	CTGAAGAGGGAGAGGGTTTGGTATCATGACGAGCCTCAGGTACGAGGGAGGACC	59248
Qy	2033	--GGCCCTACTGCGAGATGAGGGAGGCTCAGAGGACCCAGCACCTTAGGACACCG----	2087
Db	59249	CTGGCCCTCTGGGAGATGAGGAAGGCTCAGGAGACCCAGCACCCCAAGCAGGGAGCC	59308
Qy	2087	---CACCCCTGCTGAGACTGAG--GCTGCCACTTCTGGCCCTCAAGAATCAGAACATG	2140
Db	59309	CACCCCAACCCGCTGAGAAATGAGGTGCCTCTTTAGCCTCAGGAATCCAGGGATG	59366
Qy	2141	GGGACTCAGATTGCATGGGGTGGGA---CCAGGCGCTGCAAGGCTTACGGGGAGGAAGAG	2198
Db	59369	GCACTCAGGTCACGAGGGGTGGTTCCAAGCCCTCCAGGATCAGGGAAGAGAGAC	59428
Qy	2199	GAGGAGGACTCAGGGACCTTGGAAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTC	2258
Db	59429	GAGGAGGATTCAGGGGCCCTTGCAATCCAGATCAGTGGAGACCTGGGCCCTGGGAGGTC	59488
Qy	2259	CAGGGCACGGTGGCCACATATGGGCCCATATTTCTTGCTATCTTTGAGT---GACAGGAC	2314
Db	59489	CTGGCAAGGTAGCCACTGTAGTCTCATCTTCGTGCTATCTTCGAGTCTCACAGAGAGGAG	59548
Qy	2315	AGGCTCTGGTCTCAGAAGTGGGGCCCTCAGGTCACAGAGGGAGGAGTTCACGAGTCCAT	2374
Db	59549	AGGCTATGTTCTCAGGGGTGGTACTTCAGTCCGACAGGGAGGAGTCCAGGATCTAC	59608
Qy	2375	ATGGCCCAAGATGTGCC--CCCTTCATGAGGACTGGGATATCCCCGGCTCAGAAAGAGG	2433
Db	59609	AGGACCAAGGTGTCCACACTTCACGAGGAATGGGATACCTGTGGCTCAGAAAGACGG	59668
Qy	2434	GACTCCACAGTCTGGCTGTCCCTTTTAGTAGCTCTAGGGGACCATCAGGATGG	2493
Db	59669	GACCCACAGAGTCTGGCTGTCCCTGTCTTAGCTCAGGGGGACCATCAGAGGAGGATGG	59728
Qy	2494	CGGTATGTTCATTCTCAGTGTACACAGCAGGAAGTTGGGGGGCCCTCAGGGAGATG	2553
Db	59729	CCCTATGTGCCAATTCATCTGTTCCACAGCGCAGAGTGGGGAACCTTCAGGGAGATG	59788
Qy	2554	GGGTCTTGGGTAAAGGGGGATGCTACTCATGTACAGGGAATTTGGGGTTGAGGAAGCA	2613
Db	59789	AGGTTTTGGAGTAAAGGGCAATGTTTGTCTCATCTCAGGGGTTGGGGTTGAGGAAGGG	59848
Qy	2614	CAGCGCTGGCAGGAATAAGATCAGTGTAGACACACAAGGCTATTGGAATCCACACCCCA	2673
Db	59849	CAGGCCCTGTACAGAGCAAAACATGAGTACCCACAGAGGGCCATCAGAACCCCTCACCCCA	59908
Qy	2674	GAACCAAGGGGTGAGCCCTGACACCTTCACCCAGGATGTG-----GCTTCTTTTTC	2725
Db	59909	GAACCAAGGGGTGAGCCCTGGGGACCCCAACAGGGGTGACAGGATGTGGCTCTTCTC	59968
Qy	2726	ACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCTCAATCTTCAGAGGTTGACTCAGTCA	2785
Db	59969	ATTTCTGATCCAGATCTCAGTGTAGGTGAGGACCTTGTCTCAGAGGTTGACTCAGTCA	60028
Qy	2786	ACGTAGGGACCCCATCTGGTCTTAAAGACAGAGCGGTCCCAAGGATCTGCCATGCGTTGG	2845
Db	60029	CCACAGGGACCCCATCTGGTCTTACAGACAGTGTGTCCCAAGGATCTGCCAAGAGTCTCTG	60088
Qy	2846	GTGAGGAACATGAGGGAGGACTGAGGTTACCCACAGGACCAAGAACACTGAGGGACTGCA	2905
Db	60089	GTGAGGAATGTGAGGGAGGATGTAGGGTTACCAAGGGCCAGAACCGCAGATGATGACCCCA	60148
Qy	2906	CAGAAATCAGCCCTGCCCTGCTGTACCCCAAGAGCATGGGCTGGGCGGCTGTGCCGAG	2965

Db	30149	CAGAAATCAGCCCTGCCTCCTGTTGTCACCCACAGAGAGCATGGGCTTGCTTTCTTCCTGCTGAG	60208
Qy	2966	GTCTTCCTCGTTATCCTCGGATCATTGATGTACAGGACGGGAGGCCCTTGGTCTCAGAAGG	3025
Db	50209	GTCCCTCTCCTTTATCCTCGGATCACTGGTGTCACCGAGTGGGAGGCCCTTGGTCTCAGGGGG	60268
Qy	3026	CTGGCCTCAGTCACTAGAGGAGCGTCCCAGGCCCTGCCAGGAGTCAAGGTGAGGACCA	3085
Db	60269	CTGCACCCAGCTCAGTAGAGGAGGGTCCCAGGCTCTGCCAGGAGTTCAGGTGAGGACCA	60328
Qy	3086	AGCGGGCACCTCACCCAGGAGACATTAATTCCATGAATTTTGATATCTCTTGTCTGCGCCT	3145
Db	60329	AGCAGGCTCCGCATCCACGAGACATGGGTTTCCAATGAATTTTCGACATCTTTTGTGCTGCT	60388
Qy	3146	TC-CCCAAGGACCTAGGCACGTGTGGCCAGATGTTTGPCCCTCTCTGTCCTTCCATTTCT	3204
Db	60389	TCTTCGGNAGACCTTAGGCACAGGTGGCCAGATGTGGGTTTCTTAGGTTCCT---GTCCCC	60445
Qy	3205	TATCATGGATGTAACTCTTGATTGGATTTCTTCAGACCAGCAAAAGGCGCAGGATCCAGG	3264
Db	60446	TCTCAGGCATCTGAGCTCTTGATCTGAGTTTCTTCAGGCCAGCAAAAGAGTGGGATCCAGG	60505
Qy	3265	CCCTGCCAGGAAAATATAGGGCCCTCGTGAACAACAGAGGGGGTCACTCACATGTCATGA	3324
Db	60506	CCCTGCCTTGGAGAAATGTAGGGCCCTGAGTGAACACAGTAGTGGGATCATCCACATCCATGA	60565
Qy	3325	GAGTGGGATGTACACAGTCCACGCCACCCCTCGTGTAGCACGTAGAAAGCCAGGGCTGT	3384
Db	60566	GAGTGGGACCTCACAGATCCAGGCTTACCCCTCTTGATGGCACGTAGGGACCGGGGCTGT	60625
Qy	3385	GCTTTCGGCTCTGCACCTCAGGGCCCTGGATTCTCTTCTGGAGTCTCCAGGAACAGG	3444
Db	60626	GCTTACAGTCTGCACCTAAGGGCCATGATTCTCTCTTAGGAGTCTCCAGGAAACAGG	60685
Qy	3445	CAGTAGGGCTTGGTCTGAGACAGATACCTCAGGTACAGAGACAGAGATGCACAGGGTG	3504
Db	60686	CAGTAGGGCTTGGTCTGAGACAGTGTCTCAGGTTACAGAGACAGAGATGCACAGGGCTG	60745
Qy	3505	TGCCAGAGTCAATGTTTGCCCTCAATGCACACCAAGGCCCCACCTGCCACAGGACACA	3564
Db	60746	TGCCAGAGTGAATGTTTGCCCTGAATGCACCAAGGGCCCCACCTGCCACAAGACACA	60805
Qy	3565	TAGGACTCCACAGAGTCTGGCCTCACCTCCTACTGTGTCAGTCTGTAGAAATCGACCTCTG	3624
Db	60806	TAGNCTCCAAAGAGTCTGGCCTCACCTCCTACCATCAATCCTTCGAAATCGACCTCTG	60865
Qy	3625	CTGGCGGCTGTACCTGA-GTACCCCTCTCACTTCTCTCTTTCAGTTCAGGGGACAGG	3683
Db	60866	CTGGCGGCTATACCTCAGGTGCTCTCTCACTTCTCTCTTCTCAGGTTCGAGCAGACAGG	60925
Qy	3684	CCAAACAGAGGACAGGATTCCTTGGGCCACAGAGGAGCACCACAGGAGAGATCTGTA	3743
Db	60926	CCAAACG-GAGGACAGGATTCCTTGGAGGCCACAGAGGAGCACCACAGGAGAGATCTGTA	60984
Qy	3744	AGTAGGCTTTGTTAGAGTCTCCAAGTTCAAGTCTCAGCTGAGGCGCTCTCACACACTCC	3803
Db	60985	AGTAGGCTTTGTTAGAGCTCTAAGATTTGGTTCTCAGTCTGAGTCTCTCACATGCTCC	61044
Qy	3804	CTCTCTCCCCAGGCTGTGGGTCTTTCATTGGCCCAAGTCTCTGCCACACTCTCTGCTGCTG	3863
Db	61045	CTCTCTCGTAGGCTGTGGGTGCCCATTTGCCAGCTTTTGGCTGCACCTCTTGCTGCTGCTG	61104
Qy	3864	CCCTGACGAGAGTCAATCATGTCTCTTGACGAGAGGAGTCTGCATGCAAGGCTTAGGAG	3923
Db	61105	CCCTGACGAGAGTCAATCATGTCTTCTGAGCAGAGAGTTCAGCACTGCAAGGCTTAGGAG	61164
Qy	3924	CCCTTAGGCCCCAACAGAGGCGCTGGGCGCTGGTGTGTGTCAGGCTGCCAC-----	3976
Db	61165	CGTTTAGGCCCAAGAGAGGCGCTGGGCTTGTGGGTGTCACAGGCTCTCTACTGAGG	61224
Qy	3976	-----CTCCCTCTCCTCTCCTCTGGTCTGGGCAACCTTGAGGAGGTGC	4019

7	
RESULT	
RSU10688	
LOCUS	
DEFINITION	
ACCESSION	
NID	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
:	
:	
TITLE	
JOURNAL	
MEDLINE	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
FEATURES	
:	SOURCE

Db 2229 ATGAATTTTCAGATCTTTTGGCTGCTCTTCGGAAGACCTAGCACAGCTGGGCCAGATG 2288
QY 3178 TTTGTCCTCCCTCCTGCTTCCATTTCTTATCATGATGTGAACCTTTTGATTTGGATTTCT 3237
Db 2289 TGGGGTTTCTTAGTCTCT---GTTCCTCTCAGGCATGTGAGCTCTTGATCTGATTTCT 2345
QY 3238 CAGACAGCAAAAGGCGAGGATCCAGCCCTGCGAGGAAATAAAGGCGCTTCGCTGA 3297
Db 2346 CAGGCCAGCAAAAGAGTGGGATCCAGGCCCTGCGTGAGAAATGTGAGGCCCTGAGTGA 2405
QY 3298 GAACAGAGGGGTCATCCATCTGATGAGAGTGGGATGTACAGAGTCCAGCCCCACCTC 3357
Db 2406 ACACAGTGGGGATCATCCATCTCATGAGAGTGGGACCTCACAGAGTCCAGCCCTACCTC 2465
QY 3358 CTGTGACACTGAGAAGCCAGAGGCTGTGCTTTGGCTGTGACCTGAGGCCCGTGGATT 3417
Db 2466 TTGATGCACTGAGGACCGGGCTGTCTACAGTCTGACCTTAAGGCCCATGGATT 2525
QY 3418 CCTCTTCTTGAGCTCCAGGAACAGGACGTGAGGCCCTTGGTCTGAGACAGTATCCTCAG 3477
Db 2526 CCTCTCTAGGAGCTCCAGGAACAAGCAGTGAGGCCCTTGGTCTGAGACAGTGTCTCTCAG 2585
QY 3478 GTACAGAGCAGAGGATGACAGGGTGTGCCAGCAGTGAATGTTGGCTTGAATGACAC 3537
Db 2586 GTTACAGAGCAGAGGATGACAGGCTGTGCCAGCAGTGAATGTTGGCTTGAATGACAC 2645
QY 3538 CAAGGGCCCCACCTGCCACAGGACACATAGGACTTCCACAGAGTGTGCCCTCACCTCCCTA 3597
Db 2646 CAAGGGCCCCACCTGCCACAGGACACATAGGACTTCCACAGAGTGTGCCCTCACCTCCCTA 2705
QY 3598 CTGTGACTCTGAGATCGACCTCTGCTGGCCGGTGTACCCCTGA-GTACCCTCTCACT 3656
Db 2706 CCATCAATCTGCAGAAATCGACCTCTGCTGGCCGGTGTACCCCTGAGGTGCTCTCTCACT 2765
QY 3657 TCCTCTTTCAGTTTTCAGGGGACAGCCCAACCCAGAGGACAGATTCCCTGAGGCCAC 3716
Db 2766 TCCTCTTTCAGTTTTCAGGACAGAGGCCAA-CCGGAGGACAGGATTCCCTGAGGCCAC 2824
QY 3717 AGAGGACCAACAGGAGAAGATCTGAAGTAGGCTTTGTAGAGTCTCCAAAGTTTCAGT 3776
Db 2825 AGAGGACCAACAGGAGAAGATCTGAAGTAGGCTTTGTAGAGCTCTAAGATTGGT 2884
QY 3777 TCTCAGTGAAGCTCTCACACACTCCCTCTCTCCCGAGCCCTGTGGGTCTTCATTGCC 3836
Db 2885 TCTCAGTGAAGTCTCTCACATGCTCCCTCTCTCCGTAGGCCCTGTGGGTCCCATTCGCC 2944
QY 3837 AGCTCTGCCACACTCTGCTGCTGCTGCTGAGGAGTCAATGCTCTCTGAGGAGA 3896
Db 2945 AGCTTTGCTGCACTCTGCTGCTGCTGAGCAGAGTCAATGCTCTCTGAGCAGA 3904
QY 3897 GGAGTCTGCACTGCAAGCTTGAGAGCCCTTGAGGCCCAACAGAGGCCCTGGCCCTGG 3956
Db 3005 AGAGTCAAGCTGCAAGCTTGAGAGCCCTTGAGGCCCAAGAGAGGCCCTGGCCCTGG 3964
QY 3957 TGTGTGTCAGGTGGCAC-----CTCCTCTCTCTCTCTC 3992
Db 3065 TGGGTGCGCAGGCTCTACTACTGAGGAGCAGGAGGTGCTGCTCTCTCTCTCTCTC 3124
QY 3993 TGTCTTGGGACCCCTGGAGAGGTGCCACTGCTGGGTCAACAGATCTCTCCCGAGCTC 4052
Db 3125 TGTGTCCTGGACCCCTGGAGAGAGTGCCTGCTGAGTCAAGAGTCTCTCCCGAGCTC 3184
QY 4053 CTCAGGAGGCTCCGCTTTCCCACTACCAATCAACTTCACTCGACAGAGCAACCCAGTG 4112
Db 3185 CTCAGGAGGCTCTGCTTACCACTACCAATCACTTCACTTCTGCTGAGGACCAACCAATG 3244
QY 4113 AGGTTTCCAGACCGGTGAAGAGAGGGGCCAAGCACTCTTGTATCTCTGGAGTCTCTGT 4172
Db 3245 AGGTTTCCAGACCAAGAGAGAGGGGCCAAGCACTCTCTGCTGAGCAGAGTCTCTGT 3304
QY 4173 TCCAGCAGTAATCACTAAGAGGTGGCTGATTTGGTGGTTTCTGCTCTCAATATATC 4232
Db 3305 TCCGAGAGCAGCTCAGTAACAGAGGTGGATGGTTCATTTTCTGCTCCCGCAAGTATC 3364

QY 4233 GAGCCAGGAGCCAGTCACAAAGGCAAAATGCTGGAGAGTGCATCAAAAATTACAAAGC 4292
Db 3365 GAGCCAGGAGCTGGTCACAAAGGCAAAATGCTGGAGAGTGCATCAAAAATTACAAAGC 3424
QY 4293 ACTGTTTTCCTGAGATCTTTCGGAAGCCCTCTGAGTCTCTTGCAGCTGGTCTTTGGCATG 4352
Db 3425 GCTGCTTTCCTGATCTTCGCGAAAGCCCTCCGAGTCCCTGAAGATGATCTTTGGCATG 3484
QY 4353 ACCTGAAGGAGCAGACACCCACCGGCCACTCTTATGCTCTTGTACCTGCTAGGTCTCT 4412
Db 3485 ACCTGAAGGAGTGGACCCACCGCAACACCTACACCTTGTACCTGCTGGCCCTT 3544
QY 4413 CCTATGATGCTCTGCTGGGTGAATCAGATCATGCCCAAGACAGGCTTCTCTGATAATG 4472
Db 3545 CCTATGATGCTCTGCTGGGTGAATCAGATCTTCCCAAGACAGGCTTCTCTGATAATG 3604
QY 4473 TCCTGCTCATGATTGCAATGAGGCGGCCATGCTCTCTGAGGAGGAATCTGGGAGGAGC 4532
Db 3605 TCCTGGGCAATTTGCAATGGAGGCGACAGCGCTCTCTGAGGAGGAATCTGGGAGGAGC 3664
QY 4533 TGAGTGTATGAGGTGTATGATGGGAGGAGCAGTGCCTATGGGAGGCCAGGAGC 4592
Db 3665 TGGGTGTATGAGGGGTGTATGATGGGAGGAGCAGTGTCTATGGGAGGCCAGGAAAC 3724
QY 4593 TGCTCACCCAAAGATTTGGTGCAGGAAAGTACCTGGAGTA-CGGCAGGTGCCGACAGTG 4651
Db 3725 TGCTCACCCAAAGATTTGGTGCAGGAAAGTACCTGGAGTACCAGGAGTACCAGGAGTA 3784
QY 4652 ATCCCGACCTATGATGCTCTGGGGTCCAAAGGCCCTCGCTGAAACAGCATATGTA 4711
Db 3785 ATCTGCGCTATGATGCTCTGGGGTCCAAAGGCTCTGGCTGAAACAGCATATGTA 3844
QY 4712 AAGTCTTGTATGATGATCAAGTCAAGTCAAGAGTTCGCTTTTCTTCCATCCCTGC 4771
Db 3845 AAGTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3904
QY 4772 GTCAAGCAGCTTTGAGAGGAGGAGGAGGAGTCTGAGCATGATGATGATGATGATGATGAT 4831
Db 3905 GTCAAGCAGCTTTGATGAGGAGGAGGAGGAGTCTGAGCATGATGATGATGATGATGATGAT 3964
QY 4832 GTGGGAGG---GGACTGGGCCAGTGCACCTTCCAGGGCCGCTCCAGCAGTTCCTCCCT 4887
Db 3965 TGGGGAAGGGGAGGCTGGGCCAGTGCATCTAACA--GCCCTGTGCAGCAGCTTCCTT 4022
QY 4888 GCCTGCTGTGATGAGGCCCTTCTTCACTC---TGAAGAGGCGGTGAGTGTCTCA 4943
Db 4023 GCCTGCTGTAAACATGAGGCCCTTCTTCACTCTGTTGAAGAAATAGTCAGTGTCTTA 4082
QY 4944 GTAGTAGTCTTCTGTTCTATTGGTGTACTTGGAGATTATCTTTTCTCTTTTGAATT 5003
Db 4083 GTAGTGGTCTTCTATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4142
QY 5004 GTTCAATGTTTTTTTAAAGGATGTTGAATGAATCTTCAAGTCCAAAGTTCATGAATG 5063
Db 4143 GTTGAATG-TTCTTTTAAATGATGTTGAATTAACCTTCCAGCATCCAAGTTATGAATC 4201
QY 5064 ACAGCAGTCAACAGTCTGTGTATATAGTTTAAAGGTAAAGTCTTGTGTTTATTCAG 5123
Db 4202 GTAGTTAACTATGCTGTTAAATAGTTTAAAGGTAAAGTCTTGTGTTTATTCAG 4261
QY 5124 ATTGGGAATCCATTTCTTTTGTGAATTTGGG---ATAATAACAGCAGTGGAAATAGTAC 5180
Db 4262 ATTGGGAATCCCTTCTATTGTTGATTTGGACATAATAACAGCAGTGGAGTAAAGTAT 4321
QY 5181 TTGAATGTGAATAATGAGCAGTAAATAGATGAATAAGAACTAAAGAAATTAAGAG 5240
Db 4322 TTGAAGTGTG---AATTCACCGTGAATAGGTGAGAT-----AATTAAGAG 4366
QY 5241 ATAGTCAATTTCTTCCCTTATACCTCAGTCTATCTCTGTAATAATTTTAAAGATATATGCAT 5300
Db 4367 ATACTTAATTCGCGCTTATGCTCAGTCTATCTCTGTAATAATTTTAAATAATATATGCAT 4426

```

QY 5301 ACCTGGATTCTCTGGCTCTTTGAGATGTAAGAGAAATTAATCTGAATAAGAAATTC 5360
Db 4427 ACCTGGATTCTCTGGCTCTC---GTGAATGTAAGAGAAATTAATCTGAATAAATTC 4483
QY 5361 TTCCTGTTACATGGCTCTTTCTCTCCATGACATGAGCATCTGCTTTTGGAGGCGCT 5420
Db 4484 TTCTGTTAACTGGCTCATCTCTCTATGACATGAGCATCTGCTCTGTGGAAGCCCA 4543
QY 5421 GGGTTAGTAGTGAGATGCTAAGGTAAGCCAGACATCATACCACCATAGGTCGTAGAG 5480
Db 4544 GGATTAGTAGTGAGATGCTAAGGTAAGCCAGACATCATACCACCATAGGTCATTAGA 4603
QY 5481 TCTAGAGCTGCAGTCACGATATCGAGGTGGCAAGATGCTCTTAAGATGTAGGAAAA 5540
Db 4604 GTCTAGGAGCGGGTCATATAATTAAGGTGACAAGATGCTCTTAAGATGTAGGAAAA 4663
QY 5541 GTGAGAGAGGGGTAGGCTGTTGGGGTGGGCTGGGTGAGAGTGGTGTCAATGCCCTCA 5600
Db 4664 GT----ACGAGTGTGGGTATGGGCTCAGGTGAGAGTGGTGTCAATGCCCTGT 4719
QY 5601 GCTGGGGCATTTTGGGCTTTGGGAACTGCAGTTCCTCTGGGGGAGCTGATTGTAATGA 5660
Db 4720 G-TGGGGCTTTTGGCTTTGGGAACTGCATTTCTCTGAGGGATCTGATTCTAATGA 4778
QY 5661 TCTTGGGTGGATCC 5674
Db 4779 AGCTTGGTGGGTCC 4792

RESULT 8
HSU10690 HSU10690 4736 bp DNA PRI 23-JUN-1995
LOCUS Human MAGE-5b antigen (MAGE5b) gene, complete cds.
DEFINITION
ACCESSION U10690
NID 9533520
VERSION U10690.1 GI:533520
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 4736)
AUTHORS De Plaen, E., Arden, K., Traversari, C., Gaforio, J.J., Szikora, J.P.,
De Smet, C., Brasseur, F., van der Bruggen, P., Lethe, B., Lurquin, C.,
Brasseur, R., Chomez, P., De Backer, O., Cavenee, W. and Boon, T.
TITLE Structure, chromosomal localization, and expression of 12 genes of
the MAGE family
JOURNAL Immunogenetics 40 (5), 360-369 (1994)
MEDLINE 95012457
REFERENCE 2 (bases 1 to 4736)
AUTHORS De Plaen, E.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-1994) Etienne De Plaen, Ludwig Institute for
Cancer Research, 74 Avenue Hippocrate, Brussels, 1200, Belgium
FEATURES
Location/Qualifiers
1. 4736
/organism="Homo sapiens"
/isolate="patient M22"
/db_xref="taxon:9606"
/chromosome="X"
/sex="female"
/cell_type="lymphocyte"
/tissue_type="blood"
/dev_stage="adult"
644..656
/exon
2848..2934
/exon
3010..>4555
/exon
3075..3449
/gene="MAGE5b"
3075..3449
/cds

```

```

/gene="MAGE5b"
/codon_start=1
/product="MAGE-5b antigen"
/protein_id="AAA68874.1"
/db_xref="PID:9533521"
/db_xref="GI:533521"
/translation="MSLEQKSOHQKPEGLDQEBALGLVQVQATTEQEAQVSSSP
LVPQTLGKVPAAQSPGPKSPQSAIPTAIDFTLWROSIGKSSNQEEGSPSPDE
SVFRAALSKVADLIHFLLKY"
BASE COUNT 1114 a 1278 c 1304 g 1040 t
ORIGIN
Query Match 42.8%; Score 2429.6; DB 9; Length 4736;
Best Local Similarity 75.9%; Pred. No. 0;
Matches 3736; Conservative 0; Mismatches 869; Indels 315; Gaps 47;
QY 715 TTCCGCCAGGAAACATCCGGGTGCCGATGTAGAGTTCGGCCCGAAGAACTCA 834
Db 15 TTCAACCCAGGAAATCCCTGGGTGACCATGTGTGGTCCCATCTGTGCACATTTGAGT 74
QY 775 CAGAGAGAGCAGGATTTCCATTTCTGAGGAGCGGCTAGAGTTCGGCCCGAAGAACTCA 834
Db 75 CGAGAGAGCAAGGCGCTCGCTCTCAGGGGAGC-TGAGATCAGCTGAGGCGACCTGG 133
QY 835 CCCAGGCTCTGTAGGAGGCAAGGTGAGAGGTGAGGAGGAGTGTGAGGAGGAGGAGTGTG 894
Db 134 CCCTGGCTCTGTGAGGTGCAAGGTGAGATGCTGAGGAGGAGTAAAGGAGTATCCACCC 193
QY 895 CAAATAGAGAGCCCCAAATATTTCCAG---CCCCGCCCTTGTCTGCCAGCCTGGCCACCC 951
Db 194 CTGGTAGTGAGCCCCAAATATTCAGTGCACCTCTCTCTGCTAGCTCTGGACCATCC 253
QY 952 GCGGGAAGACGCTCTCAGCCTGGGCTGCCCCAGAGCCCTGCTCCAAAGAGCTTGAGAGAC 1011
Db 254 AGGGCAGGACTTCTTAGGCTGGGCCACCCAGTCCCCACCGCTTAAGCCGAGGGGA- 313
QY 1012 ACCAGGTTCTTCTCCCAAGCTCTGGAATCAGAGTGTGTGTGACAGGCGGAGGAGTGT 1071
Db 313 -----CTCAGGAGACAGAGCTTGTATGACGAGGCGAGGAGTGT 352
QY 1072 TAGGAGAGGCGAGGCGACAGGCTCTGCCAGGATCAAGATCAGCAGCCCAAGAGGAGGCG 1131
Db 353 TAGGAGAGGAGAGCTCCCA-GCTCTGCCAGGAAACAACGTCAGGAACCTTAAGGGAAGT 411
QY 1132 TGTGGGCCCCCAAGACTGCACTCCAATCCCACTCCACCCCATTCGCATTCCTCATTC 1191
Db 412 GAGGCTACCCC----- 423
QY 1192 CACCAACCCCATCTCTCTAGCTACACTCACCCCCCATCTCTCTACTCTCGTCACT 1251
Db 423 -----CACCCCAAACTCTATTCTCTGCTCCCTACCTCGTCC 457
QY 1252 CTGACCAACCCCTCCAGCCCGCAGCAGCCCAACCCCTTGTGCCACCTCACCCTCACT 1311
Db 458 CCCACTACACCCCACTTC-----CCACCCCTTCCCTACCGGACCTCTATCCACA 512
QY 1312 GCCCCCAACCCCACTCATCTCTCTATGTGCCCACTCCCATCGCTCCCTCATTCGTG 1371
Db 513 TCCCCCA-----CCCTATCTCTG 530
QY 1372 GCAGATCCGGTT-TGCCCTCTCTCAACCCAGGGAAGCCCTGTGTAGGCGCGATGTAA 1430
Db 531 GCAGATCCGATTCGCCCCCTGATTTCACCCAGGGAAGCCCTTAGGGGCGCGATGTGAT 590
QY 1431 ACCACTGACTTGAACCTCTACAGATCTGAGAGAGCCAGGTTCAATTAATGTTCTGAGGG 1490
Db 591 GCTGTGACTTGTGATTTGGGGTTCAGAGAGATCAAGG-----GCATGTTCTGAGAA 644
QY 1491 GCGGCTTGAGATCCACTGAGGAGGAGTGTGTTTAGGCTCTGTGAGGAGCAAGGTGAGTG 1550
Db 645 GCGGACTGAGATCAGCAGGAGGGAATGGCGCGGCTCTGTGAGGAGGCAAGGTGAGACC 704

```



```
FEATURES      Location/Qualifiers
source      1..2420
BASE COUNT  562 a 582 c 677 g 599 t
ORIGIN

Query Match      42.4% Score 2408; DB 5; Length 2420;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 3256 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGGCTGAGAACAGAGGGGTGATCC 3315
Db 1 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGGCTGAGAACAGAGGGGTGATCC 60

Qy 3316 ACTGCATGAGAGTGGGATGTCACAGAGTCCAGGCCACCCCTCTGGTAGCACTGAGAGC 3375
Db 61 ACTGCATGAGAGTGGGATGTCACAGAGTCCAGGCCACCCCTCTGGTAGCACTGAGAGC 120

Qy 3376 CAGGGCTGTGCTTGGGCTGCACCCCTGAGGGCCCGTGGATTCTCTTCTGGAGCTCCA 3435
Db 121 CAGGGCTGTGCTTGGGCTGCACCCCTGAGGGCCCGTGGATTCTCTTCTGGAGCTCCA 180

Qy 3436 GGAACACAGGAGTGGGCTTGGCTGACAGAGTATCTCAGGTCAACAGCAGAGGATG 3495
Db 181 GGAACACAGGAGTGGGCTTGGCTGACAGAGTATCTCAGGTCAACAGCAGAGGATG 240

Qy 3496 CACAGGCTGCCAGCAGTGAATGTTGCCCTGAATGCACACCAAGGCCCCACCTGCCA 3555
Db 241 CACAGGCTGCCAGCAGTGAATGTTGCCCTGAATGCACACCAAGGCCCCACCTGCCA 300

Qy 3556 CAGGACACATAGGACTCCACAGAGTGTGGCTCACCTCCCTACTGTCTAGTCTCTAGAA 3615
Db 301 CAGGACACATAGGACTCCACAGAGTGTGGCTCACCTCCCTACTGTCTAGTCTCTAGAA 360

Qy 3616 CGACTCTGCTGGCGGTGTACCTGAGTACCTCTCAGTCTCTCTCTCTCTCTCTCTCT 3675
Db 361 CGACTCTGCTGGCGGTGTACCTGAGTACCTCTCAGTCTCTCTCTCTCTCTCTCTCT 420

Qy 3676 GGGACAGGCCAACCCAGAGGACAGGATTCCTCGAGGCCACAGAGGACACCAAGGAGAA 3735
Db 421 GGGACAGGCCAACCCAGAGGACAGGATTCCTCGAGGCCACAGAGGACACCAAGGAGAA 480

Qy 3736 GATCTGTAAGTAGGCTTTGTTAGAGTCTCCAGGTTTCAGTCTCAGTGGGCTCTCA 3795
Db 481 GATCTGTAAGTAGGCTTTGTTAGAGTCTCCAGGTTTCAGTCTCAGTGGGCTCTCA 540

Qy 3796 CACACTCTCTCTCCAGGCTGTGGGTCTTATTCGCCAGCTCTCTGCCACACTCTCT 3855
Db 541 CACACTCTCTCTCTCCAGGCTGTGGGTCTTATTCGCCAGCTCTCTGCCACACTCTCT 600

Qy 3856 GCCTGTGCTGCTGAGGAGTCATCTCTCTTTCAGCAGAGGAGTCTGCACTGCAAGCC 3915
Db 601 GCCTGTGCTGCTGAGGAGTCATCTCTCTTTCAGCAGAGGAGTCTGCACTGCAAGCC 660

Qy 3916 TGAGGAAGCCCTTGAGGCCAACAGAGGCCCTGGGCTGTGTGTGTCAGGCTGCCAC 3975
Db 661 TGAGGAAGCCCTTGAGGCCAACAGAGGCCCTGGGCTGTGTGTGTCAGGCTGCCAC 720

Qy 3976 CTCCTCTCTCTCTCTCTGCTGCTGGGACCCCTGGAGAGGTGCCACTGCTGGGTCAAC 4035
Db 721 CTCCTCTCTCTCTCTCTGCTGCTGGGACCCCTGGAGAGGTGCCACTGCTGGGTCAAC 780

Qy 4036 AGATCTCTCCAGAGTCTCTCAGGAGGCTCGGCTCTTCCCACTACCATCTCACTCTCG 4095
Db 781 AGATCTCTCCAGAGTCTCTCAGGAGGCTCGGCTCTTCCCACTACCATCTCACTCTCG 840

Qy 4096 ACAGAGCAACCCAGTGGGTTCCAGCAGCCGTGAAGAGAGGGGCCAAGCACCCTCTTG 4155
Db 841 ACAGAGCAACCCAGTGGGTTCCAGCAGCCGTGAAGAGAGGGGCCAAGCACCCTCTTG 900

Qy 4156 TATCTGGAGTCTCTGTTCCAGCAGTAATCACTAAGAGAGTGGCTGATTTGGTGGTTT 4215
Db 1561 TATCTGGAGTCTCTGTTCCAGCAGTAATCACTAAGAGAGTGGCTGATTTGGTGGTTT 960

Qy 4216 TCTGCTCTCAAAATATCGAGCCAGGAGCCAGTCACAAAGGACGAAATGCTGGAGAGTGT 4275
Db 961 TCTGCTCTCAAAATATCGAGCCAGGAGCCAGTCACAAAGGACGAAATGCTGGAGAGTGT 1020

Qy 4276 CATCAAAAATTTACAGCACTGTTTTCTGAGATCTTCGGCAAGCCCTCTGAGTCTTGCA 4335
Db 1021 CATCAAAAATTTACAGCACTGTTTTCTGAGATCTTCGGCAAGCCCTCTGAGTCTTGCA 1080

Qy 4336 GCTGGTCTTTGGCAATTCAGCTGAAGGAGACACCCACCGGCCACTCTATGTCCTTGT 4395
Db 1381 GCTGGTCTTTGGCAATTCAGCTGAAGGAGACACCCACCGGCCACTCTATGTCCTTGT 1140

Qy 4396 CACCTGCTAGTCTCTCTCTATGATGCTCTGGGTGATTAATCAGATCATGCCCAAGAC 4455
Db 1141 CACCTGCTAGTCTCTCTCTATGATGCTCTGGGTGATTAATCAGATCATGCCCAAGAC 1200

Qy 4456 AGGCTTCTGATAATTCCTGCTCATGATTGCAATGGAGGGCGCCATGCTCTGAGGA 4515
Db 1201 AGGCTTCTGATAATTCCTGCTCATGATTGCAATGGAGGGCGCCATGCTCTGAGGA 1260

Qy 4516 GGAATCTGGAGGAGCTGAGTGTGATGGAGGTGATGATGGGAGGAGACACAGTGCCTA 4575
Db 1261 GGAATCTGGAGGAGCTGAGTGTGATGGAGGTGATGATGGGAGGAGACACAGTGCCTA 1320

Qy 4576 TGGGAGCCAGGAGGAGTCTCACCAAGATTTCGTCAGGAAAGTACCTGGAGTA-CG 4634
Db 1321 TGGGAGCCAGGAGGAGTCTCACCAAGATTTCGTCAGGAAAGTACCTGGAGTACCG 1380

Qy 4635 GCAGGTGCCGAGACAGTGCATCCGACAGCTATGATTCCTGTGGGTCCAAAGGCCCTCGC 4694
Db 1381 GCAGGTGCCGAGACAGTGCATCCGACAGCTATGATTCCTGTGGGTCCAAAGGCCCTCGC 1440

Qy 4695 TGAACACAGTATGTGAAGTCTCTGAGTATGATCAAGGTCACTGCAAGAGTTCGCTT 4754
Db 1441 TGAACACAGTATGTGAAGTCTCTGAGTATGATCAAGGTCACTGCAAGAGTTCGCTT 1500

Qy 4755 TTTCTTCCCATCCCTGCTGAAGCAGCTTTGAGAGAGGAGGAGGAGTCTGAGGATG 4814
Db 1501 TTTCTTCCCATCCCTGCTGAAGCAGCTTTGAGAGAGGAGGAGGAGTCTGAGGATG 1560

Qy 4815 AGTTGACGCAAGGCCAGTGGAGGGGAGTGGGCCAGTGCACCTTCCAGGCCCGCTCC 4874
Db 1561 AGTTGACGCAAGGCCAGTGGAGGGGAGTGGGCCAGTGCACCTTCCAGGCCCGCTCC 1620

Qy 4875 AGCAGCTTCCCTGCTGCTGACATGAGGCCCATTTCTACCTCTGAAGAGAGCGGTCA 4934
Db 1621 AGCAGCTTCCCTGCTGCTGACATGAGGCCCATTTCTACCTCTGAAGAGAGCGGTCA 1680

Qy 4935 GTGTTCTCAGTAGTGGTCTTCTGTTCTATTGGGTGACTTGGAGATTATCTTTGTTCTCT 4994
Db 1681 GTGTTCTCAGTAGTGGTCTTCTGTTCTATTGGGTGACTTGGAGATTATCTTTGTTCTCT 1740

Qy 4995 TTTGGAATTTTCAAAATGTTTTTTTTTAAGGATGTTGTAATGAATCACTCAAGTCCAAGT 5034
Db 1741 TTTGGAATTTTCAAAATGTTTTTTTTTAAGGATGTTGTAATGAATCACTCAAGTCCAAGT 1800

Qy 5055 TTATGAATGACAGCAGTCACACAGTCTGTTATATAGTTTAAAGGTAAGAGTCTTGTGT 5114
Db 1801 TTATGAATGACAGCAGTCACACAGTCTGTTATATAGTTTAAAGGTAAGAGTCTTGTGT 1860

Qy 5115 TTTATTTCAGATTGGGAATCCATTCTATTGTAATGGGTAATTAACAGCAGTGGAAAT 5174
Db 1861 TTTATTTCAGATTGGGAATCCATTCTATTGTAATGGGTAATTAACAGCAGTGGAAAT 1920

Qy 5175 AAGTACTTAGAAATGTGAATAAGTGAAGTAAATAGATGAGATAAGAACTAAAGAAAT 5234
Db 1921 AAGTACTTAGAAATGTGAATAAGTGAAGTAAATAGATGAGATAAGAACTAAAGAAAT 1980

Qy 5235 TAAGAGATAGTCAATCTGCTTATACCTCAGTCTATCTGTAATAATTTTTTAAGATAT 5294
Db 1981 TAAGAGATAGTCAATCTGCTTATACCTCAGTCTATCTGTAATAATTTTTTAAGATAT 2040
```

```
QY 5295 ATGCATACCTGGATTCCTTCCTGGCTCTTTGAGATGTAAGAGAAATTAATCTGAATAA 5354
Db 2041 ATGCATACCTGGATTCCTTCCTGGCTCTTTGAGATGTAAGAGAAATTAATCTGAATAA 2100
QY 5355 GAATTCCTCTGCTACTGCTCTTTCTCTCTCATGACATGACATCTGCTTTTGGAA 5414
Db 2101 GAATTCCTCTGCTACTGCTCTTTCTCTCTCATGACATGACATCTGCTTTTGGAA 2160
QY 5415 GGCCTGGGTAGTAGTGAGATGCTAAGGTAAAGCCAGACTATACCCACCCATAGGTC 5474
Db 2161 GGCCTGGGTAGTAGTGAGATGCTAAGGTAAAGCCAGACTATACCCACCCATAGGTC 2220
QY 5475 GTAGAGTCTAGGAGCTGACGTACGTAATCGAGGTGCAAGATGCTCTCTAAAGATGTAG 5534
Db 2221 GTAGAGTCTAGGAGCTGACGTACGTAATCGAGGTGCAAGATGCTCTCTAAAGATGTAG 2280
QY 5535 GGAAGAGTAGAGAGGGGTAGGGGTGGGGCTCCGGGTGAGAGTGGTGAGTCAATG 5594
Db 2281 GGAAGAGTAGAGAGGGGTAGGGGTGGGGCTCCGGGTGAGAGTGGTGAGTCAATG 2340
QY 5595 CCTGAGCTGGGCATTTGGGCTTTGGGAACTGCAAGTTCCTCTGCGGGAGCTGATTG 5654
Db 2341 CCTGAGCTGGGCATTTGGGCTTTGGGAACTGCAAGTTCCTCTGCGGGAGCTGATTG 2400
QY 5655 TAATGATCTTGGTGGATCC 5674
Db 2401 TAATGATCTTGGTGGATCC 2420
```

RESULT 12

```
HUMMAGIA 2420 bp DNA PRI 15-NOV-1993
LOCUS Human antigen (MAGE-1) gene, complete cds.
DEFINITION M77481
ACCESSION M77481
NID 9416114
VERSION M77481.1 GI:416114
KEYWORDS antigen.
SOURCE Homo sapiens (individual isolate patient M22) melanoma abdominal
metastasis of melanoma DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 785 to 1286)
AUTHORS van der Bruggen P., Traversari C., Chomez P., Lurquin C., De
Plaen E., Van den Eynde B., Knuth A. and Boon T.
TITLE A gene encoding an antigen recognized by cytolytic T lymphocytes on
a human melanoma
JOURNAL Science 254, 1643-1647 (1991)
MEDLINE 92086861
REFERENCE 2 (bases 1 to 2420)
AUTHORS van der Bruggen P.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-1992) Pierre van der Bruggen, Ludwig Institute
for Cancer Research, Brussels Branch, Avenue Hippocrate, 74, UCL
7459, Brussels, B-1200, Belgium
COMMENT On Nov 15, 1993 this sequence version replaced gi:187294.
FEATURES
source
1..2420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="x"
/sex="female"
/cell_line="M22-MEL.43"
1..412
/partial
/gene="MAGE-1"
/number=1
413..485
/gene="MAGE-1"
/number=2
join(413..485,561..2111)
/partial
Intron
exon
mRNA
```

```
/gene="MAGE-1"
486..560
/gene="MAGE-1"
/number=2
561..2111
/gene="MAGE-1"
/number=3
626..1555
/gene="MAGE-1"
626..1555
/gene="MAGE-1"
/codon_start=1
/product="MAGE-1"
/protein_id="AA03229.1"
/db_xref="PI:416115"
/db_xref="GI:416115"
/translation="MSLEORSLHCKPEEALQALGLVCVQAATSSSPVLVLTLE
SVPTAGSTDPPOSPOGASAFPTITRQRPSESGSREEGSPSCILESIFRAVI
TKKVADLVGLLKYRAREPVTKAEMLESVIRKHYKHFPEIFGKASESLQVFGIDVK
EADPTGHSYLVLTCLGSLYDGLDNQIMPKTGFLLIIVLMIAMEGHAPPEEIMEEL
SYMEVYDGRHSAYGEPRKLLTODLVOEKYLEYRQVPDSDPARYEFLWGPRLAETSY
VKVLEVVIKVSARVRFPPSLREALREEEGV"
polyA_signal 2095..2100
/gene="MAGE-1"
2112..2420
/partial
/gene="MAGE-1"
/number=3
BASE COUNT 562 a 582 c 677 g 599 t
ORIGIN
```

```
Query Match 42.4%; Score 2408; DB 9; Length 2420;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 3256 GGATCAGGCCCTGCCAGGAAAAATATAAGGGCCCTCGGTGAGAACAGAGGGGTCTATCC 3315
Db 1 GGATCAGGCCCTGCCAGGAAAAATATAAGGGCCCTCGGTGAGAACAGAGGGGTCTATCC 60
QY 3316 ACTGCATGAGAGTGGGATGTACAGAGTCCAGCCACCTCTCTGTAGCACTGAGAGC 3375
Db 61 ACTGCATGAGAGTGGGATGTACAGAGTCCAGCCACCTCTCTGTAGCACTGAGAGC 120
QY 3376 CAGGGCTGTCTTGGCTGTGACCTGAGGGCCGCTGATTCCTCTCTGTAGCTCCA 3435
Db 121 CAGGGCTGTCTTGGCTGTGACCTGAGGGCCGCTGATTCCTCTCTGTAGCTCCA 180
QY 3436 GGAACAGGCAGTGAAGCCCTTGGTCTGAGACAGTATCCTCAGGTACAGAGAGGATG 3495
Db 181 GGAACAGGCAGTGAAGCCCTTGGTCTGAGACAGTATCCTCAGGTACAGAGAGGATG 240
QY 3496 CACAGGGTGTGCCAGCAGTGAATGTTTGCCTGAATGCACACCAAGGCCCCACCTGCCA 3555
Db 241 CACAGGGTGTGCCAGCAGTGAATGTTTGCCTGAATGCACACCAAGGCCCCACCTGCCA 300
QY 3556 CAGGACACATAGGACTCCACAGAGTCTGGCTCACCTCCCTACTGTCTGCTGCTAGTAAT 3615
Db 301 CAGGACACATAGGACTCCACAGAGTCTGGCTCACCTCCCTACTGTCTGCTGCTAGTAAT 360
QY 3616 CGACCTCTGCTGGCCGCTGTACCTGAGTACCTCTCACTTCTCTCTCAGGTTTTCAG 3675
Db 361 CGACCTCTGCTGGCCGCTGTACCTGAGTACCTCTCACTTCTCTCTCAGGTTTTCAG 420
QY 3676 GGGACAGGCCAACCCAGAGAGGATTCCTGGAGCCACAGAGGACCAAGAGGAA 3735
Db 421 GGGACAGGCCAACCCAGAGAGGATTCCTGGAGCCACAGAGGACCAAGAGGAA 480
QY 3736 GATCTGTAAGTAGGCTTTTGTAGAGTCTCCAAGGTTCAAGTTTCTCAGCTGAGGCCCTCA 3795
Db 481 GATCTGTAAGTAGGCTTTTGTAGAGTCTCCAAGGTTCAAGTTTCTCAGCTGAGGCCCTCA 540
QY 3796 CACACTCCCTCTCTCCCGAGGCTGTGGGTCTTTCATTGTCAGCTCTCTGCCCCACACTCCT 3855
```


|||||
Db 541 CACACTCCCTCTCCCAAGGCTGTGGTCTTCAATGCCCCAGCTCCTGCCACACTCCT 600
QY 3856 GCCTGCTGCCCTGACGAGAGTCATCATGTCCTTGAGCAGAGAGTCTGCACCTGCAAGCC 3915
Db 601 GCCTGCTGCCCTGACGAGAGTCATCATGTCCTTGAGCAGAGAGTCTGCACCTGCAAGCC 660
QY 3916 TGAGGAAGCCCTTGAGGCCCAACAAGAGCCCTTGGCCCTGTGTGTGTCAGAGCTGCCAC 3975
Db 661 TGAGGAAGCCCTTGAGGCCCAACAAGAGCCCTTGGCCCTGTGTGTGTCAGAGCTGCCAC 720
QY 3976 CTCCT 4035
Db 721 CTCCT 780
QY 4036 AGATCTCTCCCAAGAGTCCTCAGGAGGCTTCCGCTCTTCCCACTACCATCAACTTCACTCG 4095
Db 781 AGATCTCTCCCAAGAGTCCTCAGGAGGCTTCCGCTCTTCCCACTACCATCAACTTCACTCG 840
QY 4096 ACAGAGGCAACCCAGTGAAGGTTCCAGCAGCCGTGAAGAGGAGGCGCCCAAGCACCTCTTG 4155
Db 841 ACAGAGGCAACCCAGTGAAGGTTCCAGCAGCCGTGAAGAGGAGGCGCCCAAGCACCTCTTG 900
QY 4156 TATCTGAGTCT 4215
Db 901 TATCTGAGTCT 960
QY 4216 TCTGTCTCTCAATATCGAGCCAGGAGGCTCAGTCAAGAGGAGGAGTCTCGAGAGTGT 4275
Db 961 TCTGTCTCTCAATATCGAGCCAGGAGGCTCAGTCAAGAGGAGGAGTCTCGAGAGTGT 1020
QY 4276 CATCAAAAATTAACAAGCACTGTTTCTCTGAGATCTTCGSCAAAGCCTCTGAGTCTTGCA 4335
Db 1021 CATCAAAAATTAACAAGCACTGTTTCTCTGAGATCTTCGSCAAAGCCTCTGAGTCTTGCA 1080
QY 4336 GCTGTCTTTGGCAGTGAAGAGCAGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 4395
Db 1081 GCTGTCTTTGGCAGTGAAGAGCAGCAGCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 1140
QY 4396 CACCTGCCTAGTCTCTCTATGATGGCTCTGGTGAATATCAGATCATGCCCAAGC 4455
Db 1141 CACCTGCCTAGTCTCTCTATGATGGCTCTGGTGAATATCAGATCATGCCCAAGC 1200
QY 4456 AGGCTTCTGTATAATGCT 4515
Db 1201 AGGCTTCTGTATAATGCT 1260
QY 4516 GGAATCTGGGAGGAGCTGATGATGAGAGTGTATGATGGGAGGAGCAGTGCCTA 4575
Db 1261 GGAATCTGGGAGGAGCTGATGATGAGAGTGTATGATGGGAGGAGCAGTGCCTA 1320
QY 4576 TGGGAGCCAGGAGCTGCTCACCAGATTTGTCAGGAGGAGTACCTGGAGTA-CG 4634
Db 1321 TGGGAGCCAGGAGCTGCTCACCAGATTTGTCAGGAGGAGTACCTGGAGTACCG 1380
QY 4635 GCAGTGGCGGACAGTATCCCGACGCTATGATGTCCTGTGGGTGCCAAGGCCCTCGC 4694
Db 1381 GCAGTGGCGGACAGTATCCCGACGCTATGATGTCCTGTGGGTGCCAAGGCCCTCGC 1440
QY 4695 TGAACACAGCTATGGAAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 4754
Db 1441 TGAACACAGCTATGGAAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 4755 TTTCTTCCATCCCTGCGTGAAGAGCTTTGAGAGGAGGAGGAGGAGTCTGAGCATG 4814
Db 1501 TTTCTTCCATCCCTGCGTGAAGAGCTTTGAGAGGAGGAGGAGGAGTCTGAGCATG 1560
QY 4815 AGTTGAGCCAAAGGCGAGTGGGAGGAGGAGTGGCCAGTGCACCTTCCAGGCCCGCTCC 4874
Db 1561 AGTTGAGCCAAAGGCGAGTGGGAGGAGGAGTGGCCAGTGCACCTTCCAGGCCCGCTCC 1620
QY 4875 AGGAGTTCCTGCTGCTGTGATGATGAGGAGGAGGAGTCTTCTACTGTGAAGAGGAGGAGTCA 4934
|||||

Db 1621 AGCAGCTTCCCTCGCTCGTGTGACATGAGGCCCATTTCTTCACTCTGAAGAGCGGTCA 1680
QY 4935 GTGTTCTCAGTAGTAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4994
Db 1681 GTGTTCTCAGTAGTAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
QY 4995 TTTGGAATTTGTTCAAAATGTTTTTTTTTAAAGGATGGTTGAATGAATCTCAGCATCCAAGT 5054
Db 1741 TTTGGAATTTGTTCAAAATGTTTTTTTTTAAAGGATGGTTGAATGAATCTCAGCATCCAAGT 1800
QY 5055 TTATGAATGACAGCAGTCACACAGTCTGTGTATATAGTTTAAAGGATGAAGAGTCTGTGT 5114
Db 1801 TTATGAATGACAGCAGTCACACAGTCTGTGTATATAGTTTAAAGGATGAAGAGTCTGTGT 1860
QY 5115 TTTATTCAGATTGGGAAATCCATCTTATTTTGTGAATTTGGGATATAACAGCAGTGGAAAT 5174
Db 1861 TTTATTCAGATTGGGAAATCCATCTTATTTTGTGAATTTGGGATATAACAGCAGTGGAAAT 1920
QY 5175 AAGTACTTAGAAATGTGAAATAGCAGTAAATAGATGAGATAAAGAACTTAAAGAAAT 5234
Db 1921 AAGTACTTAGAAATGTGAAATAGCAGTAAATAGATGAGATAAAGAACTTAAAGAAAT 1980
QY 5235 TAAGAGATAGTCAATCTTGCCTTATACCTCAGTCTATCTGTATAAATTTTAAAGATAT 5294
Db 1981 TAAGAGATAGTCAATCTTGCCTTATACCTCAGTCTATCTGTATAAATTTTAAAGATAT 2040
QY 5295 ATGCATACCTGGAATTCCTTGGCTTCTTGAAGATGTAAGAGAAATTAATCTGAATAAA 5354
Db 2041 ATGCATACCTGGAATTCCTTGGCTTCTTGAAGATGTAAGAGAAATTAATCTGAATAAA 2100
QY 5355 GAATTCCTCTGTTTCACTGGCTCTTTCTTCCATGACCTGAGCATCTGCTTTTGGAA 5414
Db 2101 GAATTCCTCTGTTTCACTGGCTCTTTCTTCCATGACCTGAGCATCTGCTTTTGGAA 2160
QY 5415 GGCCTGCTGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTATACCCACCCTAGGCTC 5474
Db 2161 GGCCTGCTGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTATACCCACCCTAGGCTC 2220
QY 5475 GTAGAGTCTAGGAGCTGAGTCACGTCACGTCAGAGTGGCAAGATGCTCTAAGATGTAG 5534
Db 2221 GTAGAGTCTAGGAGCTGAGTCACGTCACGTCAGAGTGGCAAGATGCTCTAAGATGTAG 2280
QY 5535 GGAAGAGTGAAGAGGAGTGGGTGTGGGCTCGGGTGGAGAGTGGTGGAGTGTCAATG 5594
Db 2281 GGAAGAGTGAAGAGGAGTGGGTGTGGGCTCGGGTGGAGAGTGGTGGAGTGTCAATG 2340
QY 5595 CCCTGAGCTGGGCAATTTGGCTTTGGGAACTCCAGTTCCTTCTGGGGAGCTGATG 5654
Db 2341 CCCTGAGCTGGGCAATTTGGCTTTGGGAACTCCAGTTCCTTCTGGGGAGCTGATG 2400
QY 5655 TAATGATCTTGGGTGGATCC 5674
Db 2401 TAATGATCTTGGGTGGATCC 2420

RESULT 13
HSAF002994
LOCUS HSAF002994 54155 bp DNA PRI 27-MAY-1997
DEFINITION Homo sapiens cosmids Qc4G10, Qc3C7, Qc1B11 and LK1023, complete sequence.
ACCESSION AF002994
NID G2121297
VERSION AF002994.1 GI:2121297
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 54155)
AUTHORS Gloeckner, G., Rosenthal, A., Schattevoy, R., Nordstiek, G.,
Kioschis, P., Poustka, A. and Drescher, B.
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 54155)

AUTHORS Schattevoy, R., Rosenthal, A., Drescher, B. and Schattevoy, R.
 TITLE Direct Submission
 JOURNAL Submitted (07-MAY-1997) Genome Analysis, Institut für Molekular
 Biotechnologie, Beutenbergstrasse 11, Jena 07745, Germany
 FEATURES
 Location/Qualifiers
 source
 1. 54155
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1. 38360
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="cosmid Qc4G10"
 1154..37199
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="cosmid Qc3C7"
 1833..2115
 /rpt_family="L1ME3A"
 /evidence-not_experimental
 2176..2390
 /rpt_family="L1ME3A"
 /evidence-not_experimental
 2404..2558
 /rpt_family="L1MB7"
 /evidence-not_experimental
 2573..2639
 /rpt_family="L1MB7"
 /evidence-not_experimental
 2639..2682
 /rpt_family="L1MB7"
 /evidence-not_experimental
 2699..2742
 /rpt_family="L1MB7"
 /evidence-not_experimental
 5127..5236
 /rpt_family="L1"
 /evidence-not_experimental
 5336..5434
 /rpt_family="L1"
 /evidence-not_experimental
 5435..5485
 /rpt_family="L1MB3"
 /evidence-not_experimental
 5527..5610
 /rpt_family="L1MA9"
 /evidence-not_experimental
 5611..5647
 /rpt_family="L1MA9"
 /evidence-not_experimental
 5684..5957
 /rpt_family="L1PA2"
 /evidence-not_experimental
 5983..6053
 /rpt_family="L1MA2"
 /evidence-not_experimental
 6060..6141
 /rpt_family="L1MA5"
 /evidence-not_experimental
 6145..6221
 /rpt_family="L1PA2"
 /evidence-not_experimental
 6236..6362
 /rpt_family="L1"
 /evidence-not_experimental
 6363..6592
 /rpt_family="L1MA9"
 /evidence-not_experimental
 complement(7245..7281)
 /rpt_family="ALU"
 /evidence-not_experimental
 complement(7282..7359)
 /rpt_family="ALU"
 /evidence-not_experimental

repeat_region
 complement(7396..7455)
 /rpt_family="ALU"
 /evidence-not_experimental
 complement(7456..7516)
 /rpt_family="ALU"
 /evidence-not_experimental
 7609..7691
 /rpt_family="L1ME3A"
 /evidence-not_experimental
 7840..8009
 /rpt_family="L1ME2"
 /evidence-not_experimental
 complement(8156..8198)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(8202..8263)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(8284..8306)
 /rpt_family="L1PA2"
 /evidence-not_experimental
 complement(8307..8649)
 /rpt_family="L1PA2"
 /evidence-not_experimental
 complement(8650..9050)
 /rpt_family="L1PA2"
 /evidence-not_experimental
 complement(9050..9301)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(9302..9342)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(9343..9402)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(9402..9642)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(9644..9762)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(9763..10000)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(10002..10108)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(10110..10462)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(10502..11202)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(11226..11302)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(11303..11365)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(11366..11458)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(11459..11905)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(11905..12076)
 /rpt_family="L1"
 /evidence-not_experimental
 complement(12081..12222)
 /rpt_family="L1"
 /evidence-not_experimental
 12559..14659
 /evidence-not_experimental
 misc_feature


```
|||||
Db 51772 TGGCCAGTGCACCTTCAGGGCCCCATCCATTAGCTTCCACTGCCTCGTGATATGAG 51831
Qy 4905 GCCATTCTT--CACTCTGAAGAGAGCGGTGAGTCTCTCAGTAGTAGTCTTCTGTTCTTA 4962
|||||
Db 51832 GCCATCTCCTCCTTTTGAAGAGAGCAGTCAAGATCTTACAGTGTCTGTTCTG 51891
Qy 4963 TTGGGTGACTTGGAGATTTATCTTTTGTCTCTTTTGGAAATGTTTCAAAATGTTTTTTTA 5022
|||||
Db 51892 TTGGATGACTTTGAGATTATCTTTGTTCTCTGTTGGAATGTTTCAAAATG-TTCTTTTA 51950
Qy 5023 AGGATGTTTGAATGAACCTTCAGATCAAGTTTATGAATGACAGCAGTCAAC--AGTT 5080
Db 51951 ACAAATGGTTGGATGAACCTTCAGATCAAGTTTATGAATGACAGTCAACATAGTG 52010
Qy 5081 CTGTGTATATAGTTAAAGGTAAAGAGTCTCTGTTTATTCAGATTGGGAATCCATCTCT 5140
Db 52011 CTGTTTATATAGTTAGGGTAAAGAGTCTCTGTTTATTCAGATTGGGAATCCATCTCC 52070
Qy 5141 ATTTTGTGAATG--GGATAATAACAGCAGTGGAAATAGTACTTAGAAATCT----GAAA 5194
Db 52071 ATTTTGTGAGTTGCACATAATAACAGCAGTGGAAATAGTATTTGCCCTATATTGTGAACG 52130
Qy 5195 AATGAGCAGTAAATAGATGATGATGAATAAGAACTAAAGAAATTAAGAGATAGTCAATCTTG 5254
Db 52131 AATTAGCAGTAAATATACATATACAAAGAAC-----TCAAAAGATAGTAAATCTTG 52182
Qy 5255 CTTTATACCTCAGTCTATTCTGTAAATTTTAAAGATATATGATACCTGGATTTCCTT 5314
Db 52183 CTTTATACCTCAGTCTATTATGTATAAA---TTAAATATATGTGTA-----TGTTTT 52330
Qy 5315 GCCTTTTGTGAATGTAAGAGAAATTAATCTGAATAAAGAAATTTCTTCCTGTTCACTGG 5374
Db 52231 TGCTTCTTTGAGATGCAAAAGAAATTAATCTGAATAAATTTCTTCCTGTTCACTGG 52290
Qy 5375 CTCCTTTCTTCCATGACATGACATCTGCTTTTGGAAAGCCCTGGGTTAGTAGTGA 5434
Db 52291 CTCATTTCTTTACCATCTCACCAGCATCTGCTCTGTGGAGGGCCCTGG--TAGTAGTGG 52348
Qy 5435 GATGCTAAGTGAAGCCAGCTCATACCCACCACCATAGGCTCTAGAGTCTAGGAGCTGCAG 5494
Db 52349 GATCTAAGTTAAGCCAGCTCAGCTTACCACATAGGCTCATAGAGCTAGAGCTGCAG 52408
Qy 5495 TCAGTAAATCAGGTGGCAAGATGTCCTCTTAAAGATGTAGGAAAGTGAAGAGGGGTG 5554
Db 52409 TCATGTAATTAAGGTGGCGAAGTCCCTCTAGGATGTAGTGGAAATGTAACACAGGGGTG 52468
Qy 5555 AGGCTGTGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATGCCCTGAGCTGGGGCATTTG 5614
Db 52469 AGGCTGTGGGCTCCAGGTGAGAGTGGT--GAGTATAAATGCCCTGAGCTGGGGCAATTTG 52527
Qy 5615 GCCTTTGGGAACCTGCAGTCTCTTCTGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
Db 52528 GGATCTGGGAACCTGCAGTCTCTCTGAGGAGGCTGATCTTAATGATGCCCGGTGGGTCC 52587

RESULT 14
AF030261/c 37781 bp DNA PRI 30-OCT-1997
LOCUS AF030261 Homo sapiens chromosome X clone 58F6 map Xq28, complete sequence.
DEFINITION AF030261
ACCESSION AF030261
NID 92571136
VERSION AF030261.1 GI:2571136
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 37781)
AUTHORS Gloeckner, G., Rosenthal, A., Drescher, B., Schattevoy, R., Knob, A. and
Rosenthal, A.
TITLE Sequence of cosmid 58F6 from Xq28 containing two MAGE genes
JOURNAL unpublished
```

```
2 (bases 1 to 37781)
AUTHORS Gloeckner, G., Rosenthal, A., Drescher, B., Weber, J. and Schattevoy, R.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-1997) Genome Analysis, Institut for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
Location/Qualifiers
1..37781
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="Xq28"
/clone="58F6"
1..5888
/misc_feature
/notes="overlap to AF002996"
complement(2..50)
exon
/notes="GRAIL, score = 95.000%, comment = excellent"
/evidence=not_experimental
repeat_region
complement(13..561)
/rpt_family="L1M4"
repeat_region
complement(563..685)
/rpt_family="L1M1"
exon
/evidence=not_experimental
1198..1270
/notes="GRAIL, score = 58.000%, comment = good"
/evidence=not_experimental
2065..2232
/notes="MZF, score = 59.9%"
exon
/evidence=not_experimental
repeat_region
complement(2186..2554)
/rpt_family="THE1B"
repeat_region
2851..3207
/evidence=not_experimental
/rpt_family="L1PA4"
exon
/evidence=not_experimental
complement(3187..3283)
/notes="MZF, score = 73%"
exon
/evidence=not_experimental
complement(3187..3283)
/notes="GRAIL, score = 100.000%, comment = excellent
shadow"
/evidence=not_experimental
3683..3859
/notes="GRAIL, score = 81.000%, comment = excellent"
exon
/evidence=not_experimental
3786..3859
/notes="Xpound exon prediction, score = 92% (0%)"
exon
/evidence=not_experimental
3939..4076
/notes="Xpound exon prediction, score = 87% (0%)"
exon
/evidence=not_experimental
complement(5711..5857)
/notes="GRAIL, score = 90.000%, comment = excellent"
exon
/evidence=not_experimental
5720..5949
/notes="GRAIL, score = 68.000%, comment = good shadow"
exon
/evidence=not_experimental
5722..5858
/notes="BLASTN2 (EST exons), 1:58F6_00_Ex_7 1..137 of
gbJU52239|HSU52239 Human chromosome X EST C49gl. 7/96 P =
1.8e-15 S = 483"
exon
/evidence=not_experimental
5722..5858
/notes="BLASTN2 (EST exons), 1:58F6_00_Ex_8 1..137 of
gbJU52239|HSU52239 Human chromosome X EST C49gl. 7/96 P =
1.8e-15 S = 483"
exon
/evidence=not_experimental
5761..5949
/notes="MZF, score = 56.6%"
exon
/evidence=not_experimental
6407..6458
/notes="MZF, score = 70.3%"
exon
/evidence=not_experimental
```

exon 5820..5917
/note="Xpound exon prediction, score = 72% (0%)"
/evidence=not_experimental
6916..37781
/note="overlap to AF002994"
7913..8192
/note="CpG island score = 0.86, GC = 62.10%, CpGs = 24"
/note="Region: CpG island"
/evidence=not_experimental
7914..8216
/rpt_family="AluYa5"
/evidence=not_experimental
8051..8192
/note="GC score = 11.80 (142bp)"
/note="Region: GC content"
/evidence=not_experimental
complement(9430..10349)
/note="GRAIL, score = 84.000%, comment = excellent"
/evidence=not_experimental
complement(9436..10349)
/note="Xpound exon prediction, score = 93% (0%)"
/evidence=not_experimental
complement(9457..9712)
/note="BLASTN2 (EST exons), 1:58F6_00.Ex_11 1..256 of
gb|T29724|T29724 EST92093 Homo sapiens cDNA 5' end similar
to MAGE-2 P = 8.1e-50 S = 1238"
/evidence=not_experimental
9636..10109
/note="WZEF, score = 70.5%"
/evidence=not_experimental
complement(9636..9712)
/note="BLASTN2 (EST exons), 1:58F6_00.Ex_12 1..77 of
gb|T29724|T29724 EST92093 Homo sapiens cDNA 5' end similar
to MAGE-2 P = 2.7e-11 S = 385"
/evidence=not_experimental
9636..9849
/note="GRAIL, score = 84.000%, comment = excellent shadow"
/evidence=not_experimental
complement(9637..9731)
/note="BLASTN2 (EST exons), 2:58F6_00.Ex_12 1..95 of
gb|AA378043|AA378043 EST90692 Synovial sarcoma Homo
sapiens cDNA 5' end P = 1.7e-09 S = 343"
/evidence=not_experimental
10213..10370
/note="GRAIL, score = 55.000%, comment = good shadow"
/evidence=not_experimental
complement(11315..11404)
/note="WZEF, score = 71.4%"
/evidence=not_experimental
11759..13147
/note="GC score = 43.60 (1389bp)"
/note="Region: GC content"
/evidence=not_experimental
12228..12326
/note="GRAIL, score = 74.000%, comment = good shadow"
/evidence=not_experimental
12616..12673
/note="GRAIL, score = 83.000%, comment = excellent shadow"
/evidence=not_experimental
12682..12875
/note="CpG island score = 1.00, GC = 65.50%, CpGs = 18"
/note="Region: CpG island"
/evidence=not_experimental
complement(12709..12787)
/note="GRAIL, score = 52.000%, comment = good"
/evidence=not_experimental
complement(13598..13656)
/note="WZEF, score = 55.1%"
/evidence=not_experimental
14076..14135
/note="WZEF, score = 99.9%"
/evidence=not_experimental
14076..14135

/note="BLASTN2 (EST exons), 1:58F6_00.Ex_22 1..60 of
gb|U52238|HSU52238 Human chromosome X EST C49f1.7/96 P =
0.0011 S = 231"
/evidence=not_experimental
14076..14135
/note="GRAIL, score = 100.000%, comment = excellent"
/evidence=not_experimental
complement(14141..15321)
/rpt_family="L1M4"
/evidence=not_experimental
complement(15683..15875)
/rpt_family="L1MD3"
/evidence=not_experimental
complement(15888..16252)
/rpt_family="MLT1A2"
/evidence=not_experimental
complement(16262..17531)
/rpt_family="L1MCI"
/evidence=not_experimental
17541..17701
/rpt_family="MLT1A2"
/evidence=not_experimental
17821..18010
/rpt_family="MLT1A2"
/evidence=not_experimental
complement(18071..18123)
/rpt_family="L1"
/evidence=not_experimental
18126..18361
/rpt_family="MLT2FB"
/evidence=not_experimental
18278..18377
/note="GRAIL, score = 44.000%, comment = marginal"
/evidence=not_experimental
complement(19213..19255)
/note="GRAIL, score = 61.000%, comment = good shadow"
/evidence=not_experimental
complement(19228..19766)
/rpt_family="L1M4"
Query Match 40.7%; Score 2307.6; DB 11; Length 37781;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 3479; Conservative 0; Mismatches 874; Indels 207; Gaps 44;
QY 1263 CTCCAGCCCGCCAGCAGCCAGCCAGCCCTTCTGCCACCTCAGCCTCAGCTGCCCAACCC 1322
DB 12983 CCCCCAGCCCGCCAGGATCTACAGCTCAGGATCCCGCTCCCAATCCCTACCTACAC 12924
QY 1323 CACCCCTATCTTTCATGTGCCCA--CTCCCATCGCTCCCGCCATTTGCGAGAATCC 1380
DB 12923 CAACACCATCTTCATGCTTACCCCAACCCCACTCCAGATCCCATCGGCGAGAATCC 12864
QY 1381 GGTT-TGCCCTTCTCTCAACCCAGGAGCCCTGTGTAGCCCGATGTGAAACCATGTAC 1439
DB 12863 GGTTCACCTTTCCTGCGTGAACCCAGGAGTACACG--GCCCGGATGTGACGCCATGAC 12806
QY 1440 TTGAACCTCACAGATCTGAGAGAAGCCAGGTTCAATTAATGGTCTGTAGGGGGGGTTGA 1499
DB 12805 TTGCGCATTTGGAGGTGAGGACGAGGATTC-----TCGCCCTGAGCAAGCGCTGA 12752
QY 1500 GATPCACTAGGGGAGTGGTTTGTAGGCTCTGTGAGGAGCAAGGTGAGATGTGAGGGAG 1559
DB 12751 CGTCGCGCAGGAGCAAGCAGCGCAGGCTCCGTGAGGAGCAAGGTAAAGCGCGAGGGAG 12692
QY 1560 GACTGAGGAGGACACACACCCAGGTAGATGCCCCCAAAATGATCCAGTACCACCCCTGCT 1619
DB 12691 GACTGAGGCGGCGCTCACCCAGACAGAGGGGCCCA--ATAATCCAGCTGCTCTGCT 12633
QY 1620 GCCAGCCCTGGACACCCCGCCAGGACAGATGTCTCAG-----CTGACACACCCCGG 1672
DB 12632 GCGGGGCTTGGACACCCCTCACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 12573
QY 1673 TCCGCTCCCACTGCCACTTAACCCAGAGGCAATCTGTAGTCATAGC-TTATGTGACCGG 1731

Db	12572	CCGCCACCCCGCGCGCTTTAAACGCGAGGGAACACTCTGGCGTAAGAGCTTTGTGTGACCAG	12513
Qy	1732	GGCAGGGTTGCTCAGGAGGACGAGGGCCAGGCATCAAGGTCTCAGCATCCGCCCGGCATT	1791
Db	12512	GGCAGGGCTGTTAGAGTG-----CTCAGGGCCAGACTCAGCCAGGAATC	12466
Qy	1792	AGGTCAGGACCCCTGGGAGGAACTGAGGTTCCCCACCCACACACTGTCTCTCTCATCTCC	1851
Db	12465	AAGTTCAGGACCCCAAGAGGGGACTGAGGGCAACCCACCCCTACCTCTACCTACCAATCC	12406
Qy	1852	-----ACGGCCACCCCACTCATCTCCAT	1876
Db	12405	CATCCCCAACCAACCCACCCTCATCTCTCAGAAAT-----CCCTGCTGTCAAC	12346
Qy	1877	ACCTAGCCCTTACCCCAACCTCATCTGTTCAGAAAT-----CCCTGCTGTCAAC	1925
Db	12345	TTCCCATCTCTCCCAACCACTCTCTGGCAGAAATCGGGCTTTGCCCCCTGCAATCAAC	12286
Qy	1926	CCACGGAAGCCACGGGAATGCGGCCAGGCACTCGGATCTTTGACGTCCCCATCCAGGGTC	1985
Db	12285	CCACGGAAGCTCCGGGAATGCGGCCAAGCACGCGGATCTCTGACGTTTACATGTACGG-C	12227
Qy	1986	TGATGGAGGGAAGGG-----CTTGAACAGGGCCTCAGGGGACAGAGGAG-----	2033
Db	12226	TAAGGAGGGAAGGGGTTGGGTCTCTGTGAGTATGGCTTTGGGATGCAGAGGAAGGGCCC	12167
Qy	2033	-GGCCCTACTCGAGAGTACGAGGAGGCTTCAGAGACCCAGCACCTTAGGACA-----	2084
Db	12166	AGGCCCTCTGGAAGACAGTGGAGTCTTTAGGGAGCCAGCATCCAGGACAGGGGGCCC	12107
Qy	2084	-CCGACCCCTGTCTGAGACTGAGGCTGC--CACTTTGGCCTCAAGAATCAGAACGATG	2140
Db	12106	ACTGTACCCCTGTCTCAAACCTAGCCACCTTTTCATTCAGCCGCGGAATCCTTAGGGATG	12047
Qy	2141	GGGACTCAGAT-TGCATGGGGGTGGGACCCAGGCTTGAAG-CTTAGCGGGAGAGAG	2198
Db	12046	CAGACCCACTTCAGCAGGGGGTTGGGCCACAGCCCTGCGAGGAGTCAAGGGAGGAAGAA	11987
Qy	2199	GAGGAGGACTCAGGGGACCTTGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTC	2258
Db	11986	GAGGGAGACTGAGGGGACCTTTGAGTCCAGATCAGTGTGGCAACCTTGG-CTGGGGGATC	11928
Qy	2259	CAGGGACGGTGCCACATATGCCCCATATTTCTGCTCATTTTGAAGTGAC----AGGAC	2314
Db	11927	CTGGGCACAGTGCGGAATGTGCCCGTGTCTATTGCACCTTTCAGGCTGACAGAGATTG	11868
Qy	2315	AGACTGTGCTCAGAAAGTGGGCGCTCAGTCAACAGAGAGGAGGATTCAGAGATCAT	2374
Db	11867	AGGCTGTGCTCAGGGCTGGGACTTCAGTCAAGAGAGGAGGAATCCAGAGATCTGC	11808
Qy	2375	ATGCCCAAGATGTGCCCTTTCATGAGGACTGGGATATCCCGGCTTCAGAAAGAGGG	2434
Db	11807	CGAACCAAGGTGTCCCCCTTCATGAGGACTGGGATACCCCCCGCCAGAAAGAGGG	11748
Qy	2435	ACTCCACACAGTCTGGCTGTCCCCCTTTTAGTACTGAGGGGACCAAGATCAGGGATGGC	2494
Db	11747	ATGCCACAGACTGCGCGTCTCCCTTGTCTTAGTCTGGGGGAACCTGATCAGGATGGC	11688
Qy	2495	GGTATGTTCATTTCTACTTGTACACAGGCGAAGTGTGGGGGGCCCTCAGGAGATGG	2554
Db	11687	CCTAAGTGACAAATCTCATTTGTACACAGGCGAGGTTGGGGGAACCTCAGGAGAGATAA	11628
Qy	2555	GGTCTGGGTAAAGGGGATGTCTACTCATGTACGGAAATCGGGGTTGAGGAAGCAC	2614
Db	11627	GGTGTGGTAAAGAGGAGCTGTCTGCTCAATTCAGGGGGTTGGGGTTGAGAAAGGGC	11568
Qy	2615	AGGCGCTGGCAGGAATAAGATGAGTGAGACAGACAAGGCTATTGGAATCCACACCCAG	2674
Db	11567	AGTCCCTGGCAGGAGTAAAGATGATGAACCAAGGAGGCCATCATACGTTTACCTAG	11508
Qy	2675	AACGAAGGGGTACGCCCTGGACACCTCAC-----CCAGGATGTGGCTTCTTTTC	2725

Db	11507	AACCAAAGGGGTACGCCCTGGACAAACGACACGTGGGGGTAAACAGGATGAGC-CCCTCCTC	11444
Qy	2736	ACTCCTGTTTCCAGATCTGGGCAGAGTACAGACCTCATTTCTCAGAGGTGACTCAGGTCA	2785
Db	11448	ACTTCTGTTTCCAGATCTCAGGGAGTTGATGACCTTGTGTTTTCAGAAAGGTGACTCAGGTCA	11389
Qy	2786	ACGTAGGACCCCAATCTGGTCTTAAGACAGAGCGGTCCCAAGTATCTGCCATGCGTTGCG	2845
Db	11388	ACACAGGGGCCCACTCTGTTGCGACAGATGACGTGGTGTCTAGGATCTGCCAAGCATCCAG	11329
Qy	2846	GTGAGGACATCAGGGAGGACTGAGGTACCCCAAGGACCCACAACACTGA-CGGAGACTGC	2904
Db	11328	GTGAGAGCCTTGAGGTAGGATGAGGTGAGGTACCCCTGGGCCAGAAATGCAGACAGGGGGCCCC	11269
Qy	2905	ACAGAAATCAGCCTGCCCCCTGCTGTACCCCAAGAGACATGGGTGGCGCTGTGCCGA	2964
Db	11268	ATAGAAATCTGCCTGCCCTGCGGTACTTTCAGAGACCCTGGCAGGGCTGTCACTGA	11209
Qy	2965	GGTCCTTCCTTTATCTGGGATCATTTGATGTACAGGAGCGGGAGCCCTTGCTCTGAGAAG	3024
Db	11208	AGTCCTCCATTATCTCTGGGATCTTTGATGTACAGGAAGGGAGGCCCTTGGTCTGAAGGG	11149
Qy	3025	GCTCGCTCAGGTCAAGTAGAGGGAGCGTCCACAGCCCTGCCAGAGTCAAGGTGAGGACC	3084
Db	11148	GCTGAGTCAAGGTCAAGTAGAGGGAGGGTCTCAGGCCCTGCCAGAGGTGACGTGAGGACC	11089
Qy	3085	AAGCGGGCACCTCACCCAGGACACATTAATTCGAATGAATTTGATATCTCTGCTGCC	3144
Db	11088	AAGCGACTCTCACCCAGGACACCTGGACTCCAATGAATTTGGACATCTCTCGTTGTCC	11029
Qy	3145	TTC-CCCCAAGACCTTAGCAGCGTGTGGCCAGATGTTTGTGCCCTCTGTCTCCATPCC	3203
Db	11028	TTGCGGGAGGACCTGTGTACGTATGGCCAGATGTGGTCCCTCATCTCCTTGTGTACC	10969
Qy	3204	TTATCATGGATGTGAACCTTTCG--ATTGGATTCTTCAGACCCAGAAAGGCGAGGATCC	3261
Db	10968	ATATCAGGGATGTGAGTTCTTGACATGAGAGATTCTCAAGCCAGCAAAAGGGTGGGAT-T	10910
Qy	3262	AGGCCCTGCCAGGAAAAATATAAGGGCCCTCGCTGAGAACAGAGGGGTCAATCACTGCA	3321
Db	10909	AGGCCCTACAGGAGAAAGGTGAGGGCCCTGAGTGAGCCACAGAGGGNCCCTCCACCCA	10850
Qy	3322	TGAGAGTGGGATGTACACAGAGTCCAGCCCACCCTCTGTGTAGCATGAGAACGCCAGGC	3381
Db	10849	GTAGAGTGGGACCTCACGGAGTCTGGCCAAACCCTGCTGAGACTTCTGGGAATCCGTGGC	10790
Qy	3382	TGTGCTTGGCGTCTGCACCCTGAGGGCCGTGGATTCCTCTTC-----CTGAGCTCC	3434
Db	10789	TGTGCTTGCACTCTGCACACTGAAGGCCGTGCAATTCCTCTCCAGGAATCAGAGGTCC	10730
Qy	3435	AGGAACAGGCAGTGAAGSCCTTGGTCTGAGACAGTATCCTCAGGTTCACAGAGCAGAGGAT	3494
Db	10729	AGGAACACGCAGTGAAGCCCTTGGTCTGAGTCACTGTGCTCCTCAGGTCAAGAGAGAGGGG	10670
Qy	3495	GCACAGGTTGCCAGCAGTGAATTTTGGCCCTGAATGCACACCAAGGGCCCCACCTGCC	3554
Db	10669	ACGCAGACAGTGCCTCAACACTCAAGGTTTGGCTTGAATGTCACACCAAGSGCCCCACCGCC	10610
Qy	3555	ACAGACACATAGACACTCCACAGAGTCTGGCCCTCA-CCTCCCTACTGTCACTCTGTAGA	3613
Db	10609	-CAGAACAAATGGGACTCCACAGGGCCCTGGCCTCACCTCCCTATTCTCAGTCTGTGAGC	10551
Qy	3614	ATCGACCTCTCCTGGCGGCTGTACCCTGA-GTACCCTCTCACCTCTCCTCTCAGGTTTT	3672
Db	10550	CTGAGCATGTCTGGCGGCTGTACCCTGAGGTGCCCTCCCACTTCTCTCTTCAGGTTCT	10491
Qy	3673	CA-GGGGACAGGCCCAACCAGAGACAGGATTTCCTGTAGGCCCCACAGAGGAGCACCAAGG	3731
Db	10490	GAGGGGACAGGCTGACAAGTAGGACCCGAGGCACCTGGAGGAGC-----ATTGAAGG	10439
Qy	3732	AGAGATCTGTAAGTAGGCCCTTTGTTAGAGTCTCAAG-----GTTCAAGTTCTCAGCTGA	3786
Db	10438	AGAAGATCTGTAAGTAAGCCTTTGTCAGAGGCTCCAGAGTTTCAGTTCACTGTTCTCACCTAA	10379

[illegible]

Qy	4845.	TGGGCGAGTGCACCTTCCAGGCGCGGTCCAGCAGCTTCCCTCGCTCGTGTGACATGAG	4304
Db	9298	TGGGCGAGTGCACCTTCCAGGCGCCCATCTAGTATCTCCACTGCCCTGTGTGATGAG	9239
Qy	4905	GCCATTCTT--CACATCTCAAGAGAGCGGTCAAGTCTCTCAGTAGTAGTGTTCGTGCTA	4962
Db	9328	GCCATTCCGCGCTCTTTCAGAGAGACAGTCACGATCTCTTAGCAGTGAGTTTCTGTTCTG	9179
Qy	4963	TTGGGTGACTTGGAGATTTATCTTTGTCTCTTTTGGAAATGTTCAAATGTTTTTTTTTA	5022
Db	9178	TTGGATGACTTTGAGATTTATCTTTGTTTCTCTGTTGGAAATGTTCAAATG--TTCCTTT	9120
Qy	5023	AGGGATGGTGAATGAACTTCAGCATCCAGTTTATGAATGACAGCAGTCACAC--AGTT	5080
Db	9119	ACAAATGGTGTGAATGAACTTTCAGCATCCAGTTTATGAATGACAGTAGTCACACATAGT	9060
Qy	5081	CTGTGTATATAGTTTAAAGGTAAAGAGTCTTGTGTTTTATTTCAGATTGGGAATCCATTCT	5140
Db	9059	CTGTTTATATAGTTTAAAGGTAAAGAGTCTCTGTTTTATTTCAGATTGGGAATCCATCC	9000
Qy	5141	ATTTTGTGAATTG--GGATAATAACAGCAGTGGATAAGTACTTTAGAAATGT----GAAA	5194
Db	8999	ATTTTGTGAGTTGTACATAATAACAGCAGTGGAAATGTATTTTGCCTATATTCTGAACG	8940
Qy	5195	AATGACAGCTAAATAGATGAGATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCCTG	5254
Db	8939	AATTAGCAGTAAATACATGATACAGGAAC-----TCAAAGATAGTTAATCTCTG	8888
Qy	5255	CCTTATACCTCAGTCTATTCTCTGTAATTTTTTAAAGATATATGCATACCTGGATTCCTT	5314
Db	8887	CCTTATACCTCAGTCTATTATGTGAAAA--TTAAAAATATGTGA-----TGTTTT	8840
Qy	5315	GGCTCTTTGAGAATCTAGAGAAATTAATCTGAATAAAGAAATCTTCTCTGTTCACTGG	5374
Db	8839	TGCTTCTTTGAGAATGCAAAAGAAATTAATCTGAATAAATAATCTTCTCTGTTCACTGG	8780
Qy	5375	CTCTTTTCTTCCATGCACTGAGCATCTGCTTTTGTGAAGGCCCTGGTTAGTAGTGGGA	5434
Db	8779	CTCATTTCTTTACCATCTCACTCAGCATCTGCTCTGTGGAAGGCCCTCG--TAGTAGTGGG	8722
Qy	5435	GATGCTAAGCTAAGCCAGACTCATACCCACCATAGGTCGTAGAGTCTAGGAGCTCGAG	5494
Db	8721	GATTCTAAGGTAAGCCAGACTCACGCTCACTACCATAGGCTCATAGAGTCTAGGAGCTCGAG	8662
Qy	5495	TCAGGTAATCCAGGTGGCAAGATGTCCTCTCAAGATGTTAGGGAANAAGTGAAGAGGGGTG	5554
Db	8661	TCATGTAAATTAAGGTGGCAGAAGTCCCTCTAGGTGTAGTGGAAATGTAAAGACAGGGGTG	8602
Qy	5555	AGGCTGTGGGGCTCCCGGGTGAGAGTGGTGAGTGTCAATGCCCTGAGCTGGGGCAATTTG	5614
Db	8601	AGGCTGTGGGGTTCAGAGTGAAGTGGT--GAGTATAATGCCCTGAGCTGGGGCAATTTG	8543
Qy	5615	GGCTTTGGGAAATCGCAGTTTCTCTCTGGGGGAGCTGATTGTAATGATCTCTGGGTGGATCC	5674
Db	8542	GGATCTGGGAACCTCGAGTCTCTCTCTGAAGGAGCTGATTCTTAATGATGCCCGGTGGTCC	8483

RESULT	15
LOCUS	HSAF002996
DEFINITION	Homo sapiens cosmid, complete sequence.
ACCESSION	AF002996
NID	G2121299
VERSION	AF002996.1
KEYWORDS	HTG.
SOURCE	human.
-ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 48574)
AUTHORS	Gloeckner,G., Rosenthal,A., Drescher,B., Schattevoy,R., Poust and Kioschis,P.
TITLE	Genomic Sequence around the CP1 gene

```

JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 48574)
AUTHORS Drescher, B., Rosenthal, A., Drescher, B. and Schattevoy, R.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-1997) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
source
1. .48574
/organism="Homo sapiens"
/db_xref="taxon:9606"
162. .234
/notes="GRAIL, score = 54.000%, comment = good"
/evidence-not_experimental
exon
162. .234
/notes="M2EF, score = 98.6%"
/evidence-not_experimental
exon
3730. .3924
/notes="GRAIL, score = 87.000%, comment = excellent"
/evidence-not_experimental
exon
3773. .3861
/notes="M2EF, score = 92.7%"
/evidence-not_experimental
exon
3786. .3861
/notes="Xpound exon prediction, score = 76% (0%)
/evidence-not_experimental
exon
5591. .6777
/notes="Xpound exon prediction, score = 89% (0%)
/evidence-not_experimental
exon
complement(6765..7014)
/notes="GRAIL, score = 46.000%, comment = marginal shadow"
/evidence-not_experimental
exon
6982. .7100
/notes="Alignment: EST,
AA1875771|87.4|1119|AA1574811|88.1|3103|AA3615201|96.2|
1126"
/evidence-not_experimental
exon
7010. .7100
/notes="M2EF, score = 96.8%"
/evidence-not_experimental
intron
7101. .9009
/notes="Alignment: EST, | Intron"
/evidence-not_experimental
exon
7486. .7631
/notes="Alignment: EST,
|R842671|77.6|18165|R842781|77.6|18165|AA3057141|72.3|31
95"
/evidence-not_experimental
intron
7632. .9009
/notes="Alignment: EST, | Intron"
/evidence-not_experimental
repeat_region
8297. .8464
/notes="homology = 62.50%, score = 40, counts = 84"
/rpt_type=tandem
/rpt_unit=ac
/evidence-not_experimental
repeat_region
8734. .8773
/notes="homology = 80.00%, score = 22, counts = 20"
/rpt_type=tandem
/rpt_unit=ac
/evidence-not_experimental
exon
8982. .9087
/notes="Alignment: EST,
|W006691|84.0|73175|R842671|88.5|166|242|R842781|89.7|166
|242|AA1574811|85.4|104185|AA1875771|88.8|1201199|AA30571
41189.7|961172|AA3615201|91.0|271104"
/evidence-not_experimental
exon
9010. .9087
/notes="M2EF, score = 51.3%"
/evidence-not_experimental
exon
9019. .9087
/notes="GRAIL, score = 60.000%, comment = good"
/evidence-not_experimental
intron
9088. .9166

```

```

/notes="Alignment: EST, | Intron"
/evidence-not_experimental
9167. .9294
/notes="GRAIL, score = 100.000%, comment = excellent"
/evidence-not_experimental
9167. .9229
/notes="Alignment: EST, |AA3057141|82.5|173|234"
/evidence-not_experimental
9167. .9294
/notes="M2EF, score = 97.3%"
/evidence-not_experimental
complement(9184..9392)
/notes="GRAIL, score = 45.000%, comment = marginal shadow"
/evidence-not_experimental
9217. .9294
/notes="Xpound exon prediction, score = 97% (0%)
/evidence-not_experimental
9295. .9902
/notes="Alignment: EST, | Intron"
/evidence-not_experimental
complement(9860..10093)
/notes="GRAIL, score = 42.000%, comment = marginal shadow"
/evidence-not_experimental
9903. .10013
/notes="Xpound exon prediction, score = 97% (0%)
/evidence-not_experimental
9903. .10053
/notes="GRAIL, score = 100.000%, comment = excellent"
/evidence-not_experimental
9903. .10053
/notes="Alignment: EST,
AA3615201|92.9|233|260|AA1574811|81.1|312|422|AA1875771|8
4.1|326|476|AA0456031|82.5|1136"
/evidence-not_experimental
9946. .9988
/notes="BLASTX (GENPEPT), 1. .43 of gb|M61199|HUMCS1PA_1
Human cleavage signal 1 protein mRNA, complete cds."
/evidence-not_experimental
10054. .11216
/notes="Alignment: EST, | Intron"
/evidence-not_experimental
11216. .11282
/notes="Alignment: EST,
AA0717011|89.6|5171|AA1875771|84.4|477|539|AA0456031|86.4
137|202"
/evidence-not_experimental
11283. .11917
/notes="Alignment: EST, | Intron"
/evidence-not_experimental
11918. .12382
/notes="GRAIL, score = 69.000%, comment = good"
/evidence-not_experimental
11918. .12273
/notes="Alignment: EST, |AA0717011|88.2|72|417"
/evidence-not_experimental
11918. .12382
/notes="M2EF, score = 99.4%"
/evidence-not_experimental
11918. .12045
/notes="BLASTX (GENPEPT), 59. .186 of gb|M61199|HUMCS1PA_1
Human cleavage signal 1 protein mRNA, complete cds."
/evidence-not_experimental
12198. .12385
/notes="Xpound exon prediction, score = 95% (0%)
/evidence-not_experimental
12200. .12283
/notes="TBLASTX2 (Human exons), 554. .617 of
emb|M61199|HSCS1PA Human cleavage signal 1 protein mRNA,
complete cds."
/evidence-not_experimental
12281. .12315
/notes="BLASTX (GENPEPT), 182. .216 of gb|M61199|HUMCS1PA_1
Human cleavage signal 1 protein mRNA, complete cds."

```



```

/exon          /evidence=not_experimental
12319..12381
/note="TBLASTX2 (Human exons), 741..803 of
emb|M61199|HSCS1PA Human cleavage signal 1 protein mRNA,
complete cds."
/repeat_region /evidence=not_experimental
13046..13128
/note="score = 100"
/rpt_family="ML11A"
/repeat_region /evidence=not_experimental
13209..13274
/note="score = 100"
/rpt_family="L1MC2"
/repeat_region /evidence=not_experimental
13333..13358
/note="homology = 100.00%, score = 24, counts = 13"
/rpt_type=tandem
/rpt_unit=ca
/exon          /evidence=not_experimental
complement(14962..15021)
/note="WZEF, score = 99.9%"
/exon          /evidence=not_experimental
15782..15888
/note="BLASTN (EMBL), 58..164 of emb|L18877|HSMAGE12X
Human MAGE-12 protein gene, complete cds."
/misc_feature  /evidence=not_experimental
15879..16095
/note="BLASTN (EMBL), 70..286 of emb|L18920|HSMAGE2X Human
MAGE-2 gene exons 1-4, complete cds."
/misc_feature  /evidence=not_experimental
15950..17338
/note="GC score = 42.60 (1389bp)"
/misc_feature  /evidence=not_experimental
16222..16415
/note="CpG island score = 1.00, GC = 65.50%, CpGs = 18"
/evidence=not_experimental

Query Match          40.6%; Score 2306; DB 10; Length 48574;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 3478; Conservative 0; Mismatches 875; Indels 207; Gaps 44;

QY 1263 CTCTCCAGCCCGCAGCAGCCCAAGCTTCTGCCACCTCACCTGCCCCCAACCC 1322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16114 CCCCCAGCCCGCAGGATCTACAGCTCAGGATCCCGTCCCAATCCCTACCCCTACAC 16173

QY 1323 CACCTCTATCTCTCATGTGCCCCA--CTCCCATCGCTCCCGCCCATCTTGGCAGAATCC 1380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16174 CAACACCATCTTCTATGCTTACCCCGCCAGCCATCCAGATCCCATCGGCGAGAATCC 16233

QY 1381 GGT-TGCCCTCTCTCAACCCAGGAGCCCTGTAGGCCCGATGTGNAACCATGTAC 1439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16234 GGTTCACCTTGTCCGTGAACCCAGGAGTACGG--GCCCGGATGTGACGCCACTGAC 16291

QY 1440 TTGAACCTCACAGATCTGAGGAAGCAGGTTTCATTTAATGTTCTGAGGGCGGCTTGA 1499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16292 TTGCGGATTTGGAGGTACAGGACAGCGAGATTC-----TCGCCCTTGAGCAAGCGCTGA 16345

QY 1500 GATCCACTGAGGGAGTGGTTTTAGGCTCTGTGAGGAGCAAGGTGAGATGCTGAGGGAG 1559
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16346 CGTCGCGGAGGGAAGCAGCGCAGGCTCGGTGAGGAGCAAGGTAAAGCGCGGAGGAG 16405

QY 1560 GACTGAGGAGCACACACCCAGGTAGATGGCCCCCAAAATGATCCAGTACCACCCCTGCT 1619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16406 GACTGAGGGGGCTTACCCCAACAGAGGGCCCCCA-ATAATCCAGCGCTGCTCTGCT 16464

QY 1620 GCCAGCCCTGGACACCCCGCCAGGACAGATGTCAG-----CTGGACACACCCCGG 1672
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16465 GCGGGCCCTGGACACCCCTGCGAGGGGAAGACTTCTCAGGCTGAGTCGCCACCACTCAC 16524

```

```

QY 1673 TCCTGTCCTCCATGCTCCACATTAACCCACAGGGCAATCTGTAGTCATAGC-TTATGTGACCGG 1731
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16525 CGGCCACACCCCGCCGCTTTAAACCGCAGGAACTCTGGGGTAAGAGCTTTGTGTGACCAAG 16584

QY 1732 GCAGGGTTGGTTCAGGAGGAGGAGCCAGGCATCAAGTCCAGCATCGCCCGGCATT 1791
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16585 GCGAGGGCTGGTTAGAGTG-----CTCAGGGCCAGACTCAGCCAGGAATC 16631

QY 1792 AGGGTCAGGACCTGGGAGGAACTGAGGGTTCCCCACCCACACCTCTCTCTCATCTCC 1851
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16632 AAGGTCAGGACCCCAAGAGGGGACTGAGGCAACCCACCCCTTACCTTCACTACCAATCC 16691

QY 1852 -----ACCGCCACCCCTCACTCATATCCCAT-----CCCTGCTGTCAAC 1925
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16692 CATCCCCCAACACCAACCCACCCCATCCCTCAAAACACCAACCCCAACCCCA 16751

QY 1877 ACTTACCCCTTACCCCAACCTCATCTTGTGAGAAT-----CCCTGCTGTCAAC 1925
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16752 TTCCCATCTCTCCCCACACCATCTCTGGCAGAATCCGGGCTTTGCCCTTGCATCAAC 18811

QY 1926 CCAGGAAACCCAGGGAATGGGGCCAGGCACTCTGGATCTTGACGTCCCATCCAGGGTC 1985
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16812 CCAGGAAAGCTCCGGGAATGGGGCCCAAGCAGCGGATCTGACGTTACATGTACGG-C 16870

QY 1986 TGATGAGGAAGGG-----CTTGAACAGGGCTCAGGGGAGCAGAGGAG----- 2033
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16871 TAAGGAGGAAGGGTTGGGCTCTGCTGATGATGGCTTTGGGATCAGAGGAAGGGCCC 16930

QY 2033 -GGCCCTACTCGAGATGAGGAGGCTCAGAGGACCCAGCACCTTAGGACA----- 2084
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16931 AGCCCTCTCTGGAAGACATGAGGATCTTAGGGGACCCAGCATGCCAGGAGGGGCC 16990

QY 2084 -CGCACCCCTCTCTGAGACTGAGGCTGC--CACTTCTGGCTTCAAGATCAGAAAGATG 2140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16991 ACTGTACCCCTGTCTCAACTGAGCCACCTTTTCATTCAGCCCGGGAATCCTTAGGATG 17050

QY 2141 GGGACTCAGAT-TGCATGGGGTGGAGCCAGCCCTGCAAGG-CTTACGGGAGGAGAG 2198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17051 CAGACCCACTTCAGCAGGGGGTTGGGGCCAGCCCTCGAGGAGTCAAGGGGAGGAAGAA 17110

QY 2199 GAGGAGGACTCAGGGACCTTGGATTCAGATCTGAGTCTGACCTCGGCCCTCAGAGGTC 2258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17111 GAGGAGGACTCAGGGGACCTTGGAGTCCAGATCAGTGGCACTTGG-GCTGGGGATC 17169

QY 2259 CAGGGCACGCTGGCCACATATGGCCCATATTCCTCATCTTTGAGGTGAC----AGGAC 2314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17170 CTGGGCACAGTGGCCGAATGTGCCCGTCTCATTCACCTTCAGGTTGACAGAGATTG 17229

QY 2315 AGAGCTGTGTGAGAGTGGGGCTCAGGTCAACAGAGGAGGAGTTCAGGATCCAT 2374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17230 AGGGCTGTGTGAGGCTGGGACTTCAGTTCAGCAGAGGAGGAATCCAGGATCTGC 17289

QY 2375 ATGGCCCAAGATGTGCCCTTCATGAGACTGGGGATATCCCGCTCAGAAAGAGGG 2434
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17290 CGGACCAAGGTGTGCCCTTCATGAGACTGGGATATCCCGCCCGCCCAAGAAAGGG 17349

QY 2435 ACTCCACAGACTGTGCTGTGCCCTTTTGTAGTCTTAGGGGACCAAGATCAGGATGGC 2494
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17350 ATGCCACAGACTGTGCCCTGCTCTTCTTCTAGTCTGGGGGAACCTGATCAGGATGGC 17409

QY 2495 GGTATGTTCCATCTCATTGACACAGGCAAGGTATGGGGGGCCCTCAGGAGATGG 2554
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17410 CCTAAGTGACAATCTCATTTGTACCACAGCAGGAGGTTGGGGAACCTCAGGAGATAA 17469

QY 2555 GGTCTTGGGGTAAAGGGGATGCTCTACTCATGTCAAGGNAATTGGGGTTGAGAGAGCAC 2614
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17470 GGTGTTGGTGAAGAGGAGCTGTCTCATTTCAAGGGGTTGGGGGTTGAGAAAGGGC 17529

QY 2615 AGCGCTGCGCAGGAATAAGATGAGTGAGACAGCAAGGCTATTGGAATCCACACCCAG 2674
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17530 AGTCCCTGGAAGAGTAAGATGAGTAACCCACAGGAGGCCATCATAAAGCTTCCACCTAG 17589

QY 2675 AACCAAGGGGTCAGCCCTGGACACCTCAC-----CCAGGATGTGGCTTCTTTTTC 2725

```


QY 4845 TGGGCCAGTGCACCTTCCAGGCCGCCGTCACGACGCTCCCTCGCTCGTGTGACATGAG 4904
Db 19799 TGGGCCAGTGCACCTTCCAGGCCGCCCATCCATTAGCTTCCACTGCTCGTGTGATGAG 19858
QY 4905 GCCCATTCTT--CACTCTGAACAGAGCGGTGAGTGTCTCAGTAGTAGTCTTCTTCTA 4962
Db 19859 GCCCATTCTGCTCTTTGAAGAGAGCAGTCAGCATCTTACAGTGAGTCTTCTGTCTG 19918
QY 4963 TTGGGTGACTTGCAGATTATCTTTCTCTTTTGGAAATGTTCAAAATGTTTTTTTTTA 5022
Db 19919 TTGGATGACTTTGAGATTATCTTTCTCTCTGTTGGAAATGTTCAAAAG--TTCCTTTA 19977
QY 5023 AGGGATGGTTGAATGAACCTTACGATCCAAAGTTTATGAATGACAGCAGTCACAC--AGTT 5080
Db 19978 ACAAATGGTTGGATGAACCTTACGATCCAAAGTTTATGAATGACAGTAGTCACACATAGT 20037
QY 5081 CTGTGTATATAGTTTAAAGGTAAGAGTCTGTGTTTATTACAGATTGGGAAATCCATTCT 5140
Db 20038 CTGTTTATATAGTTTAAAGGTAAGAGTCTGTGTTTATTACAGATTGGGAAATCCATTCT 20097
QY 5141 ATTTGTGAATTG--GGATAATAACAGCAGTGGATAAGTACTTAGAAATGT---GAAA 5194
Db 20098 ATTTGTGAGTTGTACATAATAACAGCAGTGGATATGTATTGCCATATGTGAACG 20157
QY 5195 AATGACGAGTAAATAGATGAGATAAGAACTAAAGAAATTAAGAGATAGTCAATCTTTG 5254
Db 20158 AATTAGCAGTAAATACATGATACAAGAAC-----TCAAAGATAGTTAATCTTTG 20209
QY 5255 CCTTATACCTCAGTCTATTCTGTAAATTTTAAAGATATATGCATACCTGGATTTCCTT 5314
Db 20210 CCTTATACCTCAGTCTATTCTGTAAATTTTAAAGATATATGCATACCTGGATTTCCTT 5374
QY 5315 GCCTTCTTTGAGATGTAGAGAAATTAATCTGAATAAGAAATCTTCTGTCTCAGTGG 5374
Db 20258 TCGTCTTTGAGATGCAAAAGAAATTAATCTGAATAATTAATCTTCTGTCTCAGTGG 20317
QY 5375 CTCCTTTCTTCCATGCACTGAGCATCTGCTTTTGGAGGCCCTGGGTTAGTAGTGA 5434
Db 20318 CTCATTTCTTTACCATTCACTCAGCATCTGCTGTGGAAGGCCCTGG--TAGTAGTGG 20375
QY 5435 GATGCTAAGGTAAAGCCAGACTCATACCCACCCTAGGTCGTAGAGCTAGGAGCTGCAG 5494
Db 20376 GATTTAAGGTAAAGCCAGACTCAGCTCTACCCATAGGTCATAGAGCTAGGAGCTGCAG 20435
QY 5495 TCACGTAAATCGAGGTGGCAAGATGCTCTTAAAGATGTAGGAAAGAGTCAGAGAGGGGTG 5554
Db 20436 TCATGTAAATTAAGGTGGCAGAGTCTCTAGATGTAGTGAATGTAAAGACAGGGGTG 20495
QY 5555 AGGGTGTGGGGTCCGGGTGAGAGTGGTGGAGTGAATGCCCTGAGCTGGGGCATTTTG 5614
Db 20496 AGGGTGTGGGGTCCAGAGTGGAGTGGT--GAGTATAAATGCCCTGAGCTGGGGCAATTTG 20554
QY 5615 GCCTTTGGGAACTGCAGTCTCTTCTGGGGGAGCTGATTGTAATGATCTTGGTGGATCC 5674
Db 20555 GGATCTGGGAACCTGCAGTCTCTTCTGAAGGAGCTGAATCTAATGATGCCGGTGGGTC 20614

Search completed: December 10, 1999, 07:37:09
Job time: 18018 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 1999, 02:53:33 ; Search time 189.19 Seconds
(without alignments)
7503.514 Million cell updates/sec

Title: US-08-819-669D-8
Perfect score: 5674
Sequence: 1 CCGGGGACCACTGGCATC.....TAATGATCTGGTGGGATCC 5674

Scoring table: IDENTITY_NUC

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5674	100.0	5674	1	Tumour rejection a
2	5672.4	100.0	5674	1	MAGE-1 nucleic aci
3	5650.8	99.6	5724	1	Tumour rejection a
4	2419	42.6	2419	1	Antigen E gene. Nu
5	2419	42.6	2419	1	Tumour rejection a
6	2415.8	42.6	2419	1	M22-MEL antigen E
7	2408	42.4	2420	1	Tumour rejection a
8	2408	42.4	2420	1	Human melanoma ant
9	2058.8	36.3	4157	1	Tumour rejection a
10	2057.2	36.3	4157	1	MAGE-2 gene. Nucle
11	1762.6	31.1	4204	1	CDNA encoding MAGE
12	1678.8	29.6	2531	1	Tumour rejection a
13	1678.8	29.6	2531	1	Tumour rejection a
14	1677.2	29.6	2531	1	MAGE-4 gene. Nucle
15	1675.6	29.5	2531	1	MAGE-41 gene. Nucl
16	1541	27.2	1691	1	Tumour rejection a
17	1504.6	26.5	2305	1	MAGE-51 genomic DN
18	1504.6	26.5	2305	1	Tumour rejection a
19	1504.6	26.5	2305	1	MAGE-51 gene. Dete
20	1410.8	24.9	2226	1	MAGE-5 cDNA. Nucle
21	1410.8	24.9	2226	1	Tumour rejection a
22	988.8	17.4	1640	1	Tumour rejection a
23	976	17.2	1640	1	MAGE-3 cDNA. Nucle
24	900.2	15.9	1810	1	MAGE-8 genomic DNA
25	900.2	15.9	1810	1	Tumour rejection a
26	900.2	15.9	1810	1	MAGE-8 gene. Deter
27	875.2	15.4	1375	1	MAGE-6 cDNA. Tumou
28	834.4	14.7	1094	1	H6/MAGE-1 expressi
29	832.8	14.7	1084	1	H6/MAGE-1 expressi
30	830.6	14.6	1947	1	MAGE-7 genomic DNA
31	830.6	14.6	1947	1	Tumour rejection a
32	830.6	14.6	1947	1	MAGE-7 gene. Deter
33	815.2	14.4	1412	1	Tumour rejection a
34	815.2	14.4	1412	1	MAGE-9 gene. Deter
35	812	14.3	1412	1	MAGE-9 genomic DNA
36	760	13.4	1022	1	MAGE-4 encoding ge
37	750.8	13.2	1068	1	Tumour rejection a
38	749.2	13.2	1068	1	MAGE-4 cDNA. Nucle
39	583.4	10.3	1107	1	MAGE-11 genomic DN
40	583.4	10.3	1107	1	Tumour rejection a
41	583.4	10.3	1107	1	MAGE-11 gene. Dete
42	540.6	9.5	943	1	MAGE-31 gene. Nucl
43	540.6	9.5	943	1	Tumour rejection a

ALIGNMENTS

RESULT 1
Q72477

ID Q72477 standard; DNA; 5674 BP.

AC Q72477; (first entry)

DE Tumour rejection antigen MAGE-1 encoding DNA.

KW Tumour rejection antigen; melanoma antigen-1; MAGE-1; MAGE-3;

KW cancer; cytolytic T cells; antigen D; human leucocyte antigen;

OS Homo sapiens.

FN Key Location/Qualifiers

FT cds 3881..4711

FT /*tag= a

PN W09423031-A.

RD 13-OCT-1994.

PF 17-MAR-1994; U02877

PR 26-MAR-1993; US-037230.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Boon-falleur T, Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;

DR WPI; 94-333192/41.

PT New tumour rejection antigen precursor MAGE3 - useful in

PS treatment and diagnosis of cancer

CC Example 26; Page 59; 105pp; English.

CC Q72477 is the DNA sequence which encodes melanoma antigen-1

CC (MAGE-1). Another melanoma antigen MAGE-3 is encoded by Q72470,

CC this is a tumour rejection antigen precursor. Melanomas

CC characterised by the expression of MAGE-3 can be detected, or

CC monitored, by contacting a test sample with an agent that can

CC recognise MAGE-3. The melanoma can be treated by the administration

CC of cytolytic T cells specific for the complex of antigen D (the

CC mature rejection antigen derived from MAGE-3) and a human leucocyte

CC antigen (esp. HLA-A1).

SQ Sequence 5674 BP; 1276 A; 1644 C; 1569 G; 1185 T;

Query Match 100.0%; Score 5674; DB 1; Length 5674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGGGGGACCACTGGCATCCCTCCCTTACACCCCAATCCCTCCCTTTAGCCACCC 60

Db 1 CCGGGGGACCACTGGCATCCCTCCCTTACACCCCAATCCCTCCCTTTAGCCACCC 60

QY 61 ATCCAAACATCTTCACGCTCACGCCAGCCAGCCAGCAATCCGGTCCACCCCTG 120

Db 61 ATCCAAACATCTTCACGCTCACGCCAGCCAGCCAGCAATCCGGTCCACCCCTG 120

QY 121 CTCTCAACCACTGAGAGCCAGGTCCAGAGTGTGACGCCACTGACTTGTAGGATG 180

Db 121 CTCTCAACCACTGAGAGCCAGGTCCAGAGTGTGACGCCACTGACTTGTAGGATG 180

QY 181 TTAGAGAGAGCGAGGTTTCGGTCTGAGGGCGCGCTTGGATCGGTGAGGAGCGGG 240

Db 181 TTAGAGAGAGCGAGGTTTCGGTCTGAGGGCGCGCTTGGATCGGTGAGGAGCGGG 240

QY 241 CCCAGCTCTGTAAAGAGCGAAGTGCATGCTGAGGAGGAGTGGAGCCACTTACCCC 300

Db 241 CCCAGCTCTGTAAAGAGCGAAGTGCATGCTGAGGAGGAGTGGAGCCACTTACCCC 300

QY 301 AGATAGAGACCCCAATAATCCCTTCATGCGAGTCCCTGACCATCTGGTGTGACTTC 360

Db 301 AGATAGAGACCCCAATAATCCCTTCATGCGAGTCCCTGACCATCTGGTGTGACTTC 360

QY 361 TCAGGCTGGGCGACCCCGAGCCCTTGTGCTTAAACCACTGGGACTCGAAGTCAGAG 420

Db 361 TCAGGCTGGGCGACCCCGAGCCCTTGTGCTTAAACCACTGGGACTCGAAGTCAGAG 420

44 448.516 7.9 2559 1 V26618 MAGE-10 tumour rej
45 384.8 6.8 2940 1 V69727 Tumour rejection a

|||||
Db 2581 ACTCATGTTCAGGAATGGGGTTGAGGAAGCACAGCGCTGGCAGAAATANAAGATCAGT 2640
|||||
QY 2641 GAGACAGACAGGCTATTGGAATCCACACCCAGAACCAAGGGGTGAGCCCTGGACACC 2700
|||||
Db 2641 GAGACAGACAGGCTATTGGAATCCACACCCAGAACCAAGGGGTGAGCCCTGGACACC 2700
|||||
QY 2701 TCACCCAGGATGTGGGTTCTTTTTCACCTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCT 2760
Db 2701 TCACCCAGGATGTGGGTTCTTTTTCACCTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCT 2760
|||||
QY 2761 CATCTCAGAGGTGACTCAGGTACAGGTAGGAGACCCCAATCTGGTCTAAAGACAGAGCG 2820
Db 2761 CATCTCAGAGGTGACTCAGGTACAGGTAGGAGACCCCAATCTGGTCTAAAGACAGAGCG 2820
|||||
QY 2821 GTCCAGGATCTGCCATGCTTCGGGTGAGGAACATGAGGAGGACTGAGGGTACCCACG 2880
Db 2821 GTCCAGGATCTGCCATGCTTCGGGTGAGGAACATGAGGAGGACTGAGGGTACCCACG 2880
|||||
QY 2881 GACCAGAACTGAGGAGACTGCACAGAAATCAGCCCTGCCCTGCTGTCAACCCACAG 2940
Db 2881 GACCAGAACTGAGGAGACTGCACAGAAATCAGCCCTGCCCTGCTGTCAACCCACAG 2940
|||||
QY 2941 AGCATGGGTGGGCGCTGCTGAGAGGCTGCGCTCAGGTACGTAGAGGAGGCTGCCAGGCC 3000
Db 2941 AGCATGGGTGGGCGCTGCTGAGAGGCTGCGCTCAGGTACGTAGAGGAGGCTGCCAGGCC 3000
|||||
QY 3001 ACGGGAGGCTTGGTCTGAGAGGCTGCGCTCAGGTACGTAGAGGAGGCTGCCAGGCC 3060
Db 3001 ACGGGAGGCTTGGTCTGAGAGGCTGCGCTCAGGTACGTAGAGGAGGCTGCCAGGCC 3060
|||||
QY 3061 CTGCCAGGAGTCAAGGTGAGGACCAAGCGGACCTCACCAGGACACATTAATTCAT 3120
Db 3061 CTGCCAGGAGTCAAGGTGAGGACCAAGCGGACCTCACCAGGACACATTAATTCAT 3120
|||||
QY 3121 GAATTTTGATATCTTGCTGCTGCTCCCAAGGACCTAGGACAGTGTGGCCAGATGTTT 3180
Db 3121 GAATTTTGATATCTTGCTGCTGCTCCCAAGGACCTAGGACAGTGTGGCCAGATGTTT 3180
|||||
QY 3181 GTCCCTCTGCTCTTCATCTTATCATGGATGTGAACCTCTTGATTTGGATTTCTCAG 3240
Db 3181 GTCCCTCTGCTCTTCATCTTATCATGGATGTGAACCTCTTGATTTGGATTTCTCAG 3240
|||||
QY 3241 ACCAGCAAAAGGACGAGTCCAGGCTTCCAGGAAATATTAAGGCGCTGCGTGAGAA 3300
Db 3241 ACCAGCAAAAGGACGAGTCCAGGCTTCCAGGAAATATTAAGGCGCTGCGTGAGAA 3300
|||||
QY 3301 CAGAGGGGTATCCACTGATGAGAGTGGGGATGTACAGAGTCCAGCCACCCCTCTG 3360
Db 3301 CAGAGGGGTATCCACTGATGAGAGTGGGGATGTACAGAGTCCAGCCACCCCTCTG 3360
|||||
QY 3361 GTAGCACTGAGAACCCAGGCTGTGCTTGGGCTGTCACACCTCAGGGCCGCTGGATTCCT 3420
Db 3361 GTAGCACTGAGAACCCAGGCTGTGCTTGGGCTGTCACACCTCAGGGCCGCTGGATTCCT 3420
|||||
QY 3421 CTTCTCGAGCTCCAGGAACCCAGGCTGAGGCTTGGTCTGAGACAGTATCTCTCAGGTC 3480
Db 3421 CTTCTCGAGCTCCAGGAACCCAGGCTGAGGCTTGGTCTGAGACAGTATCTCTCAGGTC 3480
|||||
QY 3481 ACAGACAGAGGATGACAGAGGTGTGCCAGCAGTGAATTTGCCCTGAATGACACACAA 3540
Db 3481 ACAGACAGAGGATGACAGAGGTGTGCCAGCAGTGAATTTGCCCTGAATGACACACAA 3540
|||||
QY 3541 GGCCCCACTGCCACAGGACATAGGACTCCACAGAGTCTGCCCTCAGCTCCCTACTG 3600
Db 3541 GGCCCCACTGCCACAGGACATAGGACTCCACAGAGTCTGCCCTCAGCTCCCTACTG 3600
|||||
QY 3601 TCAGTCTGTAGAAATGACCTCTGCTGGCCGGCTGTACCTGAGTACCCCTCACTTCCT 3660
Db 3601 TCAGTCTGTAGAAATGACCTCTGCTGGCCGGCTGTACCTGAGTACCCCTCACTTCCT 3660
|||||
QY 3661 CTTTCAAGGTTTCAGGGACAGGCCAACCCAGAGGACAGGATTCCTTGGAGGCCACAGAG 3720
|||||

Db 3661 CTTTCAAGGTTTCAGGGACAGGCCAACCCAGAGGACAGGATTCCTTGGAGGCCACAGAG 3720
QY 3721 GAGCAACAAGGAGAAGATCTGTAAGTAGGCTTTGTTAGAGTCTCCAAGGTTTCAAGTTCTC 3780
|||||
Db 3721 GAGCAACAAGGAGAAGATCTGTAAGTAGGCTTTGTTAGAGTCTCCAAGGTTTCAAGTTCTC 3780
|||||
QY 3781 AGCTGAGGCTCTCACACACTCCTCTCTCCAGGCTGTGGGTCTTCAATGCCAGCT 3840
Db 3781 AGCTGAGGCTCTCACACACTCCTCTCTCCAGGCTGTGGGTCTTCAATGCCAGCT 3840
|||||
QY 3841 CTTGCCACACTCCTGCCCTGCTGCTGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTGCTG 3900
Db 3841 CTTGCCACACTCCTGCCCTGCTGCTGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTGCTG 3900
|||||
QY 3901 TCTGCACTTCAAGCCTTGAAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTGCTG 3960
Db 3901 TCTGCACTTCAAGCCTTGAAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTGCTG 3960
|||||
QY 3961 TGTGAGGCTTCCAGCCT 4020
Db 3961 TGTGAGGCTTCCAGCCT 4020
|||||
QY 4021 CACTGCTGGGTCAACAGATCCTCCAGAGTCTCTCAGGAGGCTTCCAGGCTTCCAGGCTTCC 4080
Db 4021 CACTGCTGGGTCAACAGATCCTCCAGAGTCTCTCAGGAGGCTTCCAGGCTTCCAGGCTTCC 4080
|||||
QY 4081 CATCAACTTCACTCGACAGAGGCAAGCCTGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4140
Db 4081 CATCAACTTCACTCGACAGAGGCAAGCCTGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4140
|||||
QY 4141 GCGAAGCCTCTTGTATCTCTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4200
Db 4141 GCGAAGCCTCTTGTATCTCTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4200
|||||
QY 4201 TGATTTGGTGGTCTGCTCTCTCAAAATACAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4260
Db 4201 TGATTTGGTGGTCTGCTCTCTCAAAATACAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4260
|||||
QY 4261 AATGCTGGAGAGTGTCTCAAAATACAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4320
Db 4261 AATGCTGGAGAGTGTCTCAAAATACAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4320
|||||
QY 4321 CTTGAGTCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4380
Db 4321 CTTGAGTCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4380
|||||
QY 4381 CTTCTATGCTTGTCTCAGTCT 4440
Db 4381 CTTCTATGCTTGTCTCAGTCT 4440
|||||
QY 4441 GATCATGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4500
Db 4441 GATCATGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4500
|||||
QY 4501 CCATGCTCTGAGGAGAAATCTGGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4560
Db 4501 CCATGCTCTGAGGAGAAATCTGGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4560
|||||
QY 4561 GGAGCAGAGTCTTGGGAGCCAGGAGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4620
Db 4561 GGAGCAGAGTCTTGGGAGCCAGGAGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4620
|||||
QY 4621 GTACCTGGAGTACGGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4680
Db 4621 GTACCTGGAGTACGGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4680
|||||
QY 4681 CCAAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4740
Db 4681 CCAAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC 4740
|||||
QY 4741 GCAAGAGTTCGCTTTTCTTCCATCCCTCGTGGTGAAGCAGCTTTCAGAGAGGAGGAGAG 4800
Db 4741 GCAAGAGTTCGCTTTTCTTCCATCCCTCGTGGTGAAGCAGCTTTCAGAGAGGAGGAGAGAG 4800
|||||

```
QY 4801 GGAGTCTGAGCATGAGTTGACCCAGGCCAGTGGGAGGGGAGTGGCCAGTGCACCTT 4860
DB 4801 GGAGTCTGAGCATGAGTTGACCCAGGCCAGTGGGAGGGGAGTGGCCAGTGCACCTT 4860
QY 4861 CCAGGGCCGCTCCAGCAGCTTCCCTCGCTGCTGACATGAGGCCCATCTTCACTCT 4920
DB 4861 CCAGGGCCGCTCCAGCAGCTTCCCTCGCTGCTGACATGAGGCCCATCTTCACTCT 4920
QY 4921 GAAGAGAGGGTCACTGTTCTCAGTAGTAGGTTCTGTTCTATTCGCTCACTTGGAGATT 4980
DB 4921 GAAGAGAGGGTCACTGTTCTCAGTAGTAGGTTCTGTTCTATTCGCTCACTTGGAGATT 4980
QY 4981 TATCTTTGTTCTTTTGGAAATGTTTCAAAATGTTTTTAAAGGATGGTGAATGAAC 5040
DB 4981 TATCTTTGTTCTTTTGGAAATGTTTCAAAATGTTTTTAAAGGATGGTGAATGAAC 5040
QY 5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAAGG 5100
DB 5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAAGG 5100
QY 5101 TAAAGAGTCTGTGTTTATTCAGATTGGAAATCCATTCTATTTTGTGAATTTGGGATAT 5160
DB 5101 TAAAGAGTCTGTGTTTATTCAGATTGGAAATCCATTCTATTTTGTGAATTTGGGATAT 5160
QY 5161 AACAGCAGTGGAAATAGTACTTAGAAATGTGAAATGAGCAGTAAATAGATGAGATAA 5220
DB 5161 AACAGCAGTGGAAATAGTACTTAGAAATGTGAAATGAGCAGTAAATAGATGAGATAA 5220
QY 5221 AGAATCAAGAAATTAAGAGATAGTCAATTTCTTGCTTATACCTCAGTCTATTTCTGTAAA 5280
DB 5221 AGAATCAAGAAATTAAGAGATAGTCAATTTCTTGCTTATACCTCAGTCTATTTCTGTAAA 5280
QY 5281 ATTTTAAAGATATATGATACCTGAGTTTCTTGGCTTCTTTGAGAAATGTAAGAGAAAT 5340
DB 5281 ATTTTAAAGATATATGATACCTGAGTTTCTTGGCTTCTTTGAGAAATGTAAGAGAAAT 5340
QY 5341 TAAATCTGAATAAAGAAATTTCTTCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCA 5400
DB 5341 TAAATCTGAATAAAGAAATTTCTTCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCA 5400
QY 5401 TCTGCTTTTGAAGGCCCTGGGTTAGTAGTGGAGATGCTTAAGGTAAGCCAGACTCATAC 5460
DB 5401 TCTGCTTTTGAAGGCCCTGGGTTAGTAGTGGAGATGCTTAAGGTAAGCCAGACTCATAC 5460
QY 5461 CCACCCATAGGTCGTAGAGTCTAGAGTCGAGTCAGTCAGTCAATTCAGGTGGCAAGATGTC 5520
DB 5461 CCACCCATAGGTCGTAGAGTCTAGAGTCGAGTCAGTCAATTCAGGTGGCAAGATGTC 5520
QY 5521 CTCTAAAGATGTAGGAAAAGTGAAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580
DB 5521 CTCTAAAGATGTAGGAAAAGTGAAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580
QY 5581 GTGGAGTCTAATGCCCTGAGCTGGGGCATTTTGGGCTTTTGGGAACTGCGAGTTCTCTTCT 5640
DB 5581 GTGGAGTCTAATGCCCTGAGCTGGGGCATTTTGGGCTTTTGGGAACTGCGAGTTCTCTTCT 5640
QY 5641 GGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
DB 5641 GGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
```

RESULT 2

```
Q32352
ID Q32352 standard; DNA: 5674 BP.
AC Q32352:
DT 22-APR-1993 (first entry)
DE MAGE-1 nucleic acid.
KW melanoma antigen; MAGE TRA; melanoma antigen tumor rejection antigen;
KW tumor rejection antigen precursor; MAGE; antigen E; gene family; ss.
OS Homo sapiens.
FH key Location/Qualifiers
FT *cds 3881..4711
```

/*tag= a

```
FT WO9220356-A.
PD 26-NOV-1992.
PE 22-MAY-1992; U04354.
PR 23-MAY-1991; US-705702.
PR 09-JUL-1991; US-728838.
PR 23-SEP-1991; US-764364.
PR 12-DEC-1991; US-807043.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;
PI Van Den Eynde B, Van Der Bruggen P, Van Pel A;
DR WPI; 92-415450/30.
PT Nucleic acid mol.: encoding a human tumour rejection antigen
PT precursor - useful as an immunostimulant in a vaccine for
PT treating and preventing cancers, also useful in diagnosis
PS Disclosure; Page 71-73, 142pp; English.
CC The sequences given in Q32352-69 represent a new family of genes
CC referred to as melanoma antigens (MAGE). The cDNAs of this gene
CC family were identified during the isolation of the antigen E gene.
CC The MAGE cDNAs, when tested, did not transfer expression of antigen
CC E, but they did show substantial homology to the antigen E cDNA
CC sequence. The MAGE DNAs share a certain degree of homology with each
CC other and are expressed in tumour cells including several types of
CC human tumour cells as well as in human tumors. MAGE expression is not
CC restricted to melanomas. MAGE refers to a family of tumor rejection
CC antigen precursors. The antigens resulting from these genes are
CC referred to as MAGE TRAs or melanoma antigen tumor rejection antigens.
CC See also Q32351.
SQ Sequence 5674 BP; 1277 A; 1644 C; 1568 G; 1185 T;
```

Query Match 100.0%; Score 5672.4; DB 1; Length 5674;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 CCCGGGGCACCACCTGGCATCCCTCCCTACACGCCCAATCCCTCTTACGCCACCC 60
DB 1 CCCGGGGCACCACCTGGCATCCCTCCCTACACGCCCAATCCCTCTTACGCCACCC 60
QY 61 ATCCAAACATCTTACCGTCTACCCGCCAGCCCAAGCCAGGAGATCCGGTTCACCCCTG 120
DB 61 ATCCAAACATCTTACCGTCTACCCGCCAGCCCAAGCCAGGAGATCCGGTTCACCCCTG 120
QY 121 CTCTCAACCCAGGAGAACCCAGGTGCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180
DB 121 CTCTCAACCCAGGAGAACCCAGGTGCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180
QY 181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGCGGCTTGAGATCGGTGGAGGAAGCGGG 240
DB 181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGCGGCTTGAGATCGGTGGAGGAAGCGGG 240
QY 241 CCCAGCTCTGTAAAGAGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300
DB 241 CCCAGCTCTGTAAAGAGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300
QY 301 AGATAGAGGACCCCAATAATCCCTTTCATGCGAGTCCCTGGACCATCTGCTGGTGACATC 360
DB 301 AGATAGAGGACCCCAATAATCCCTTTCATGCGAGTCCCTGGACCATCTGCTGGTGACATC 360
QY 361 TCAGGCTGGGCCACCCCGAGCCCTTGTCTCTTAAACACTGGGGAGCTCGAAGTCAGAG 420
DB 361 TCAGGCTGGGCCACCCCGAGCCCTTGTCTCTTAAACACTGGGGAGCTCGAAGTCAGAG 420
QY 421 CTCGCTGTGATCAGGGAAGGCTGCTTAGGAGAGGGCAGCGCTCCAGGCTCTGCCAGACAT 480
DB 421 CTCGCTGTGATCAGGGAAGGCTGCTTAGGAGAGGGCAGCGCTCCAGGCTCTGCCAGACAT 480
QY 481 CATGCTCAGGATTTCTCAAGGAGGCTGAGGCTCCCTTAAGACCCCACTCCCGTGACCCCAAC 540
DB 481 CATGCTCAGGATTTCTCAAGGAGGCTGAGGCTCCCTTAAGACCCCACTCCCGTGACCCCAAC 540
QY 541 CCCCACTCCAATGCTCACTCCCGTGACCCCACTCTTTCATTTGTCATTCACACCCCA 600
DB 541 CCCCACTCCAATGCTCACTCCCGTGACCCCACTCTTTCATTTGTCATTCACACCCCA 600
```


Db 541 CCCCCCTCAATGCTCACTCCCGTGACCCAAACCCCTCTTCAATTGTATTCACACCCCA 600
Qy 601 CCCCCATCCCCCACCACCCCTCCCTCAACCTGATGCCATCCGCCAGCCATTCACCCCT 660
Db 601 CCCCCATCCCCCACCACCCCTCCCTCAACCTGATGCCATCCGCCAGCCATTCACCCCT 660
Qy 661 CACCCCCACCCACCCACCCACCCACCCCTCCCTCAACCTGATGCCATCCGCCAGCCATTCACCCCT 720
Db 661 CACCCCCACCCACCCACCCACCCACCCCTCCCTCAACCTGATGCCATCCGCCAGCCATTCACCCCT 720
Qy 721 CAGGAAACATCCGGGTGCCGGATGTGAGCCCACTGACTTGGCATTGTGGGGCAGAGA 780
Db 721 CAGGAAACATCCGGGTGCCGGATGTGAGCCCACTGACTTGGCATTGTGGGGCAGAGA 780
Qy 781 GAAGCAGGTTTCCATTCTGAGGAGCGGCTAGAGTTGGCCCAAGAACTCAACCCAGG 840
Db 781 GAAGCAGGTTTCCATTCTGAGGAGCGGCTAGAGTTGGCCCAAGAACTCAACCCAGG 840
Qy 841 CTCGTGAGGAGCAAGGTGAGAGGCTGAGGAGGACTGAGGACCCGCCACTTCCAATA 900
Db 841 CTCGTGAGGAGCAAGGTGAGAGGCTGAGGAGGACTGAGGACCCGCCACTTCCAATA 900
Qy 901 GAGAGCCCAATAATTCAGAGCCCGCCCTTGTGCGCAGCCCTGGCCACCCCGCGGAAGA 960
Db 901 GAGAGCCCAATAATTCAGAGCCCGCCCTTGTGCGCAGCCCTGGCCACCCCGCGGAAGA 960
Qy 961 CGTCTCAGCCTGGGCTGCCCGCAGACCCCTGCTTCCAAAGCCCTTGAGAGACACAGGTT 1020
Db 961 CGTCTCAGCCTGGGCTGCCCGCAGACCCCTGCTTCCAAAGCCCTTGAGAGACACAGGTT 1020
Qy 1021 TTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACAGGCGAGGACTGGTTAGGAGAG 1080
Db 1021 TTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACAGGCGAGGACTGGTTAGGAGAG 1080
Qy 1081 GCAGGCGCAGGCTCTGCCAGGATCAAGATCAGCACCACCCAGAGGAGGCTGTGGGCC 1140
Db 1081 GCAGGCGCAGGCTCTGCCAGGATCAAGATCAGCACCACCCAGAGGAGGCTGTGGGCC 1140
Qy 1141 CCAAGACTGCACCTCAATCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCA 1200
Db 1141 CCAAGACTGCACCTCAATCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCA 1200
Qy 1201 CCAATCTCTCAGCTACACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCA 1260
Db 1201 CCAATCTCTCAGCTACACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCA 1260
Qy 1261 ACCCTCAGGCCCCAGCAGCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCA 1320
Db 1261 ACCCTCAGGCCCCAGCAGCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCA 1320
Qy 1321 CCAACCTCTCATCTCTCATGTGCCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCC 1380
Db 1321 CCAACCTCTCATCTCTCATGTGCCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCC 1380
Qy 1381 GGTTCGCCCTGTCTCAACCCAGGAGCCCTGGTAGGCGCGATGTGAACCACTGACT 1440
Db 1381 GGTTCGCCCTGTCTCAACCCAGGAGCCCTGGTAGGCGCGATGTGAACCACTGACT 1440
Qy 1441 TGAACCTCAGAGATCTGAGAGAGCCAGGTTCAATTTAAGGTTCTGAGGGCGGCTTGAG 1500
Db 1441 TGAACCTCAGAGATCTGAGAGAGCCAGGTTCAATTTAAGGTTCTGAGGGCGGCTTGAG 1500
Qy 1501 ATCCACTGAGGAGGTGGTTTGGGCTGTGAGGAGGCAAGGTGAGATGCTCAGGAGG 1560
Db 1501 ATCCACTGAGGAGGTGGTTTGGGCTGTGAGGAGGCAAGGTGAGATGCTCAGGAGG 1560
Qy 1561 ACTGAGGAGCACACACCCAGGATAGTGGCCCAAAATGATCCAGTACCACCCCTGCTG 1620
Db 1561 ACTGAGGAGCACACACCCAGGATAGTGGCCCAAAATGATCCAGTACCACCCCTGCTG 1620
Qy 1621 CAGCCCTTGACACCCCGGCGCAGAGATGTCTCAGCTGGACACCCCGGCTCCGCTCC 1680
Db 1621 CAGCCCTTGACACCCCGGCGCAGAGATGTCTCAGCTGGACACCCCGGCTCCGCTCC 1680

Qy 1681 CACTGCCACTTAACCCACAGGCAATCTGTAGTCATAGCTTATGTACGGGGCAGGGT 1740
Db 1681 CACTGCCACTTAACCCACAGGCAATCTGTAGTCATAGCTTATGTACGGGGCAGGGT 1740
Qy 1741 GGTCAAGAGAGGAGGCGCCAGGCATCAAGTCCAGCATCCGCCCGGCATTTAGGTCAG 1800
Db 1741 GGTCAAGAGAGGAGGCGCCAGGCATCAAGTCCAGCATCCGCCCGGCATTTAGGTCAG 1800
Qy 1801 ACCCTGGAGGAACTGAGGTTCCCCACCCACACCTGTCTCTCATCTCCACCCGACC 1860
Db 1801 ACCCTGGAGGAACTGAGGTTCCCCACCCACACCTGTCTCTCATCTCCACCCGACC 1860
Qy 1861 CCACTCACATTTCCCATACCTACCCCTACCCCAACCTCATCTTGTCAAGATCCCTGCTG 1920
Db 1861 CCACTCACATTTCCCATACCTACCCCTACCCCAACCTCATCTTGTCAAGATCCCTGCTG 1920
Qy 1921 TCAACCCAGGAGCCACCGGAATGGCGCCAGGCATCGGATCTTGAGCTCCCATCCA 1980
Db 1921 TCAACCCAGGAGCCACCGGAATGGCGCCAGGCATCGGATCTTGAGCTCCCATCCA 1980
Qy 1981 GGTCTGTAGTGGAGGAGGCTTGACAGGCGCTCAGGGGAGCAGAGGAGGCGCTTAC 2040
Db 1981 GGTCTGTAGTGGAGGAGGCTTGACAGGCGCTCAGGGGAGCAGAGGAGGCGCTTAC 2040
Qy 2041 TCGAGATGAGGAGGCGCTCAGAGGACCCAGACACCTAGACACCCCTGTCTGAG 2100
Db 2041 TCGAGATGAGGAGGCGCTCAGAGGACCCAGACACCTAGACACCCCTGTCTGAG 2100
Qy 2101 ACTGAGGCTGCCACTTCTGGCCTCAAGATCAGAACGATGGGACTCAGATTCATGGGG 2160
Db 2101 ACTGAGGCTGCCACTTCTGGCCTCAAGATCAGAACGATGGGACTCAGATTCATGGGG 2160
Qy 2161 GTGGGACCCAGGCTCCAGGCTTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
Db 2161 GTGGGACCCAGGCTCCAGGCTTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
Qy 2221 GGAATCCAGATCAGTGTGAGGCTCGGCCCTGAGAGGTCCAGGCGCGGTCACACATATG 2280
Db 2221 GGAATCCAGATCAGTGTGAGGCTCGGCCCTGAGAGGTCCAGGCGCGGTCACACATATG 2280
Qy 2281 GGCATATTTCTGCACTTTTGTAGGTGACAGGACAGAGGTGTGGTCTGAGAAGTGGGGCC 2340
Db 2281 GGCATATTTCTGCACTTTTGTAGGTGACAGGACAGAGGTGTGGTCTGAGAAGTGGGGCC 2340
Qy 2341 TCAGGTCAACAGAGGAGGAGTTCAGGATCCATATGGCCCAAGATGTGCCCTTCATG 2400
Db 2341 TCAGGTCAACAGAGGAGGAGTTCAGGATCCATATGGCCCAAGATGTGCCCTTCATG 2400
Qy 2401 AGGACTGGGGATATCCCGGCTCAGAAAGAGGACTCCACACAGTCTGGCTGTCCCTT 2460
Db 2401 AGGACTGGGGATATCCCGGCTCAGAAAGAGGACTCCACACAGTCTGGCTGTCCCTT 2460
Qy 2461 TTAGTAGCTCTAGGGGACAGATCAGGATGGCGGTATTTCCATTCCTACTTGTACCA 2520
Db 2461 TTAGTAGCTCTAGGGGACAGATCAGGATGGCGGTATTTCCATTCCTACTTGTACCA 2520
Qy 2521 CAGGAGGAGTGGGGGCGCTCAGGAGATGGGGTCTTGGGTAAAGGGGGATGTCT 2580
Db 2521 CAGGAGGAGTGGGGGCGCTCAGGAGATGGGGTCTTGGGTAAAGGGGGATGTCT 2580
Qy 2581 ACTCATGTCAAGGAATTTGGGGTTGAGGAGCAGAGCGCTGGCAGGAATAAAGATGAGT 2640
Db 2581 ACTCATGTCAAGGAATTTGGGGTTGAGGAGCAGAGCGCTGGCAGGAATAAAGATGAGT 2640
Qy 2641 GAGACAGACAAGCTATTGGAATCCACACCCAGAACCCAGAGGGGTGAGCCCTGGACACC 2700
Db 2641 GAGACAGACAAGCTATTGGAATCCACACCCAGAACCCAGAGGGGTGAGCCCTGGACACC 2700
Qy 2701 TCACCCAGATGTGGCTTCTTTTCTACTCTCTTTCAGATCTGGGGCAGGTGAGGACCT 2760
Db 2701 TCACCCAGATGTGGCTTCTTTTCTACTCTCTTTCAGATCTGGGGCAGGTGAGGACCT 2760

Db 4921 GAAGACAGCGGTCACTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGTGACTTGAGATT 4980
Qy 4981 TATCTTTGTTCTTTTGGAAATGTTTCAAAATGTTTTTTTAAAGGATGGTTGAATGAAC 5040
Db 4981 TATCTTTGTTCTTTTGGAAATGTTTCAAAATGTTTTTTTAAAGGATGGTTGAATGAAC 5040
Qy 5041 TTCACATCCAAAGTTTATGAATGACAGCAGTCACACAGTCTGTTGATATAGATTAAAGG 5100
Db 5041 TTCACATCCAAAGTTTATGAATGACAGCAGTCACACAGTCTGTTGATATAGATTAAAGG 5100
Qy 5101 TAAGAGTCTGTTGTTTATTTCAGATTGGAAATCAATCTATTGTTGAAATGGGATAT 5160
Db 5101 TAAGAGTCTGTTGTTTATTTCAGATTGGAAATCAATCTATTGTTGAAATGGGATAT 5160
Qy 5161 AACACAGTGGAAATAGTACTTTAGAAATGTGAAATGAGCAGTAAATAGATGAGATAA 5220
Db 5161 AACACAGTGGAAATAGTACTTTAGAAATGTGAAATGAGCAGTAAATAGATGAGATAA 5220
Qy 5221 AGAATCAAGAAATTAAGAGATAGTCAATCTTGGCTTATACCTCAGTCTATTCGTAAA 5280
Db 5221 AGAATCAAGAAATTAAGAGATAGTCAATCTTGGCTTATACCTCAGTCTATTCGTAAA 5280
Qy 5281 ATTTTAAAGATATATGATACCTCGATTTCCTTGCTTCTTTGAGATGTAAGAGAAAT 5340
Db 5281 ATTTTAAAGATATATGATACCTCGATTTCCTTGCTTCTTTGAGATGTAAGAGAAAT 5340
Qy 5341 TAAATCTGAATAAAGAAATCTTCTGTTCACTGGCTCTTTCTTCTCCATGCACTGAGCA 5400
Db 5341 TAAATCTGAATAAAGAAATCTTCTGTTCACTGGCTCTTTCTTCTCCATGCACTGAGCA 5400
Qy 5401 TCTGTTTTTGAAGCGCCCTGGGTTAGTAGTGAGATGCTAAGGTAAGCAGACTCATAC 5460
Db 5401 TCTGTTTTTGAAGCGCCCTGGGTTAGTAGTGAGATGCTAAGGTAAGCAGACTCATAC 5460
Qy 5461 CCACCATAGGTCGTAGAGTCTAGAGCTGAGAGCTGCAGTCAGTCACTGAGTGGCAAGATGC 5520
Db 5461 CCACCATAGGTCGTAGAGTCTAGAGCTGAGAGCTGCAGTCACTGAGTGGCAAGATGC 5520
Qy 5521 CTCTAAAGATGATAGGAAAGTGAAGAGGGGTGAGGTTGCGGCTCCGGGTGAGAGTG 5580
Db 5521 CTCTAAAGATGATAGGAAAGTGAAGAGGGGTGAGGTTGCGGCTCCGGGTGAGAGTG 5580
Qy 5581 GTGGAGTGTCAATGCCCTGAGCTGGGGCAATTTTGGCTTTGGGAAACTGCAGTCTCCTCT 5640
Db 5581 GTGGAGTGTCAATGCCCTGAGCTGGGGCAATTTTGGCTTTGGGAAACTGCAGTCTCCTCT 5640
Qy 5641 GGGGAGCTGATGTTGAATGATCTTGGTGGATCC 5674
Db 5641 GGGGAGCTGATGTTGAATGATCTTGGTGGATCC 5674

RESULT 3

Q98902
ID Q98902 standard; DNA; 5724 BP.
AC Q98902:
DE 28-FEB-1996 (first entry)
DT Tumour rejection antigen (MAGE-1) gene.
KW Tumour rejection antigen; MAGE-1; monoclonal antibody; Mab;
KW diagnosis; immunoassay; cancer; ss.
OS Homo sapiens.
FH key Location/Qualifiers
FT cds 3881..4711
FT /*tag= a
FT /product= Tumour rejection antigen MAGE-1.
FT /note= "The CDS is not indicated in the text of the
FT specification but is suggested in the layout
FT of the sequence."
PN W09520974-A1.
PD 10-AUG-1995.
PF 05-JAN-1995; U00095.
PR 01-FEB-1994; US-190411.
PA (LUDW-) LUDWIG INST CANCER RES.
PA (SLOK) SLOAN KETTERING INST CANCER RES.

PA (SLOK) MEMORIAL SLOAN-KETTERING CANCER CENT.
PI Boon-fallour T, Chen Y, Garin-chesa P, Old LJ, Rettig WJ;
PI Stockert E, Van der bruggen P;
DR WPI: 95-283606/37.
PT New monoclonal antibody binding specifically to MAGE-1 - useful for
PT diagnosis and monitoring of cancer, also new hybridomas, recombinant
PT MAGE-1 and immunogenic peptide(s)
PS Disclosure: Page 16-19; 33pp; English.
CC A monoclonal antibody directed against the tumour rejection antigen
CC (MAGE-1) can be used to detect MAGE-1 in samples by standard
CC immunassay methods for diagnosis and monitoring of cancer etc. The
CC monoclonal antibody is designated MA454 and is produced by the
CC hybridoma deposited as ATCC HB11540. The monoclonal antibody is
CC specific for MAGE-1, having no reactivity for MAGE-2 or MAGE-3.
CC Peptide fragments of MAGE-1 (See R80618-20) may be useful as
CC immunogens for production of the monoclonal antibody and antisera.
SQ Sequence 5724 BP; 1282 A; 1653 C; 1589 G; 1200 T;

Query Match 99.6%; Score 5650.8; DB 1; Length 5724;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5652; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCGGGGCACCACCTGGCATCCCTCCCTTACCACCCCAATCCCTCCCTTTAGCCACCC 50
Db 1 CCCGGGGCACCACCTGGCATCCCTCCCTTACCACCCCAATCCCTCCCTTTAGCCACCC 50
Qy 61 ATCCAAACATCTTCACGCTCACCCCAAGCCAGCAGCAATCCGGTTCACACCCCTG 120
Db 61 ATCCAAACATCTTCACGCTCACCCCAAGCCAGCAGCAATCCGGTTCACACCCCTG 120
Qy 121 CTCTCAACCCAGGGAAGCCAGTCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180
Db 121 CTCTCAACCCAGGGAAGCCAGTCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180
Qy 181 TTAGAGAGAGCGAGGTTTTCGGTCTGAGGGCGGCTTGAGATCGGTGAGGAAGCGGG 240
Db 181 TTAGAGAGAGCGAGGTTTTCGGTCTGAGGGCGGCTTGAGATCGGTGAGGAAGCGGG 240
Qy 241 CCACGCTCTGAAGAGGCAAGGTGACATGCTGAGGAGGACTGAGGACCCACTTACCCC 300
Db 241 CCACGCTCTGAAGAGGCAAGGTGACATGCTGAGGAGGACTGAGGACCCACTTACCCC 300
Qy 301 AGATAGAGACCCCAATAATCCCTTCATGCCAGTCTTGACCACATCTGGTGGGACTTC 360
Db 301 AGATAGAGACCCCAATAATCCCTTCATGCCAGTCTTGACCACATCTGGTGGGACTTC 360
Qy 361 TCAGGCTGGGCGACCCCGACGCCCTTGTGCTTAAACCACTGGGACTCGAAGTCAGAG 420
Db 361 TCAGGCTGGGCGACCCCGACGCCCTTGTGCTTAAACCACTGGGACTCGAAGTCAGAG 420
Qy 421 CTCCTGTGATCAGGAAGGGGCTGCTTAGGAGAGGCGACGCTCCAGGCTCTGCCAGACAT 480
Db 421 CTCCTGTGATCAGGAAGGGGCTGCTTAGGAGAGGCGACGCTCCAGGCTCTGCCAGACAT 480
Qy 481 CATGCTCAGGATTCCTCAAGAGGGGCTGAGGGTCCCTTAAGACCCCACTCCCGTACCACAC 540
Db 481 CATGCTCAGGATTCCTCAAGAGGGGCTGAGGGTCCCTTAAGACCCCACTCCCGTACCACAC 540
Qy 541 CCCCACTCAATGCTCACTCCCGTACCCCAACCCCTCTTCAATGTCATTCACACCCCA 600
Db 541 CCCCACTCAATGCTCACTCCCGTACCCCAACCCCTCTTCAATGTCATTCACACCCCA 600
Qy 601 CCCCACTCCCAACCCCACTCCCTCAACCTGATGCCATCCGCGGAGGATTCACACCT 660
Db 601 CCCCACTCCCAACCCCACTCCCTCAACCTGATGCCATCCGCGGAGGATTCACACCT 660
Qy 661 CACCCCGACCCCGACCCCGACCCCACTCCCAACCCCGACCCCGAGGATTCGGTTCCCG 720
Db 661 CACCCCGACCCCGACCCCGACCCCACTCCCAACCCCGAGGAGGATTCGGTTCCCG 720
Qy 721 CCAGGAACATCCGGTGCCCGATGTGACGCCACTGACTTGGCATTGTGGGCGAGAGA 780
Db 721 CCAGGAACATCCGGTGCCCGATGTGACGCCACTGACTTGGCATTGTGGGCGAGAGA 780

Db 721 CCAGGAACATCCGGGTCCCGGATGTACGCCACTGACTTGCOCATTTGTGGGCAGAGA 780
Qy 781 GAACCGAGGTTTCCATTCTAGGACGCGGTAGAGTTTCGGCCGAAGAACTGTACCCAGG 840
Db 781 GAACCGAGGTTTCCATTCTAGGACGCGGTAGAGTTTCGGCCGAAGAACTGTACCCAGG 840
Qy 841 CTCGTGTAGGAGCAAGGTGAGAGGCTGAGGAGTGTAGGACCCGCGCCACTCAAATA 900
Db 841 CTCGTGTAGGAGCAAGGTGAGAGGCTGAGGAGTGTAGGACCCGCGCCACTCAAATA 900
Qy 901 GAGAGCCCCAAATATTCCAGCCCCGCGCTTGTCTGCCAGCCCTGCCCCACCCGCGGAGA 960
Db 901 GAGAGCCCCAAATATTCCAGCCCCGCGCTTGTCTGCCAGCCCTGCCCCACCCGCGGAGA 960
Qy 961 CGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAGCCTTTGAGAGACACCAGGTT 1020
Db 961 CGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAGCCTTTGAGAGACACCAGGTT 1020
Qy 1021 TTTCCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACGAGGACGAGTGTAGGAGG 1080
Db 1021 TTTCCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACGAGGACGAGTGTAGGAGG 1080
Qy 1081 GCAGGACAGGCTCTGCCAGGCATCAAGATCAGACCCCAAGAGGAGGCTGTGGGCC 1140
Db 1081 GCAGGACAGGCTCTGCCAGGCATCAAGATCAGACCCCAAGAGGAGGCTGTGGGCC 1140
Qy 1141 CCAAGACTGCACCTCCAATCCCACTCCACCCCATTCGCATTCCCATTTCCCAACCAACC 1200
Db 1141 CCAAGACTGCACCTCCAATCCCACTCCACCCCATTCGCATTCCCATTTCCCAACCAACC 1200
Qy 1201 CCCATCTCTCAGTACACTCTCACCCCATTCCTACTCTCTACTCTCTACTCTCTACTCT 1260
Db 1201 CCCATCTCTCAGTACACTCTCACCCCATTCCTACTCTCTACTCTCTACTCTCTACTCT 1260
Qy 1261 ACCCTCAGCCCCAGCACCCAGCCCAACCTTCTGCCACCTCACTCTCACTGCCCCAAC 1320
Db 1261 ACCCTCAGCCCCAGCACCCAGCCCAACCTTCTGCCACCTCACTCTCACTGCCCCAAC 1320
Qy 1321 CCCACCTCATCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTTGGCAGAAATCC 1380
Db 1321 CCCACCTCATCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTTGGCAGAAATCC 1380
Qy 1381 GGTTCGCCCTGTCTCAACCCAGGAGGACCCCTGGTAGGCCGATGTAAACCACTGACT 1440
Db 1381 GGTTCGCCCTGTCTCAACCCAGGAGGACCCCTGGTAGGCCGATGTAAACCACTGACT 1440
Qy 1441 TGAACCTCAGAGATCTGAGAAAGCCAGGTTCAATTAATGTTCTGAGGGCGCTTGAG 1500
Db 1441 TGAACCTCAGAGATCTGAGAAAGCCAGGTTCAATTAATGTTCTGAGGGCGCTTGAG 1500
Qy 1501 ATCCACTGAGGGAGTGTGTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAG 1560
Db 1501 ATCCACTGAGGGAGTGTGTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAG 1560
Qy 1561 ACTGAGGAGCACACACCCAGGTAGATGGCCCCAAATGATCAGTACCAACCCCTGCTG 1620
Db 1561 ACTGAGGAGCACACACCCAGGTAGATGGCCCCAAATGATCAGTACCAACCCCTGCTG 1620
Qy 1621 CCAGCCCTGGACACCCGCGCAGACAGATGCTCAGCTGGACACCCCGCTCCGCTCC 1680
Db 1621 CCAGCCCTGGACACCCGCGCAGACAGATGCTCAGCTGGACACCCCGCTCCGCTCC 1680
Qy 1681 CACTGCCACTTAACCCACAGGCAATCTGTAGTCATAGCTTATGTACCCGGGCGAGGTT 1740
Db 1681 CACTGCCACTTAACCCACAGGCAATCTGTAGTCATAGCTTATGTACCCGGGCGAGGTT 1740
Qy 1741 GGTGAGGAGGACGAGGCCCCAGGCATCAAGGTCCAGCATCGCCCCGCAATTAGGGTCAGG 1800
Db 1741 GGTGAGGAGGACGAGGCCCCAGGCATCAAGGTCCAGCATCGCCCCGCAATTAGGGTCAGG 1800
Qy 1801 ACCCTGGGAGGAACTGAGGTTTCCCAACCCACACCTGTCTCTCATCTCCACCCGACCC 1860
Db 1801 ACCCTGGGAGGAACTGAGGTTTCCCAACCCACACCTGTCTCTCATCTCCACCCGACCC 1860

Qy 1861 CCACCTACATTTCCATACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 1920
Db 1861 CCACCTACATTTCCATACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 1920
Qy 1921 TCACCCACCGAAGCCACGGGAATGCGGCGCAGGCACTCGGATTTGACGTGCCCATCCA 1980
Db 1921 TCACCCACCGAAGCCACGGGAATGCGGCGCAGGCACTCGGATTTGACGTGCCCATCCA 1980
Qy 1981 GGGTCTCATGAGGGAAGGGCTTGAACAGGGCTTGAACAGGGCTTGAACAGGGCTTGAAC 2040
Db 1981 GGGTCTCATGAGGGAAGGGCTTGAACAGGGCTTGAACAGGGCTTGAACAGGGCTTGAAC 2040
Qy 2041 TGGCAGATGAGGAGGCGCTCAGAGGACCCAGCACCTTAGACACCCGACCCCTCTCTGAG 2100
Db 2041 TGGCAGATGAGGAGGCGCTCAGAGGACCCAGCACCTTAGACACCCGACCCCTCTCTGAG 2100
Qy 2101 ACTGAGGCTGCCACTTCTGSCCTCAAGAAATCAGAAATCAGAAATCAGAAATGAGGACT 2160
Db 2101 ACTGAGGCTGCCACTTCTGSCCTCAAGAAATCAGAAATCAGAAATGAGGACTTGCATGGG 2160
Qy 2161 GTGGGACCCAGGCTGCAAGGCTTACGCGGAGGAAGAGGAGGAGGACTCAGGGGACCTT 2220
Db 2161 GTGGGACCCAGGCTGCAAGGCTTACGCGGAGGAAGAGGAGGAGGACTCAGGGGACCTT 2220
Qy 2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGCTGGCCACATATG 2280
Db 2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGCTGGCCACATATG 2280
Qy 2281 GCCCATATTTCTGCTCATCTTTGAGGTGACAGGACAGAGTGTGCTGTGAGAGTGGGGCC 2340
Db 2281 GCCCATATTTCTGCTCATCTTTGAGGTGACAGGACAGAGTGTGCTGTGAGAGTGGGGCC 2340
Qy 2341 TCAGGTCAACAGAGGAGGAGTTCAGAGTCCATATGCCCCAAGATGTGCCCTTCTCATG 2400
Db 2341 TCAGGTCAACAGAGGAGGAGTTCAGAGTCCATATGCCCCAAGATGTGCCCTTCTCATG 2400
Qy 2401 AGGACTGGGATATCCCGGCTCAGAAAGAGGACTTCCACAGTCTGGCTGTGCCCTT 2460
Db 2401 AGGACTGGGATATCCCGGCTCAGAAAGAGGACTTCCACAGTCTGGCTGTGCCCTT 2460
Qy 2461 TTAGTAGCTCTAGGGGACAGATCAGGATGGGGTATGTTCCATTTCTCATTGTACCA 2520
Db 2461 TTAGTAGCTCTAGGGGACAGATCAGGATGGGGTATGTTCCATTTCTCATTGTACCA 2520
Qy 2521 CAGCAGGAAGTTGGGGGCGCTCAGGAGATGGGTCTTGGGTAAGGGGGATGCT 2580
Db 2521 CAGCAGGAAGTTGGGGGCGCTCAGGAGATGGGTCTTGGGTAAGGGGGATGCT 2580
Qy 2581 ACTCATGTCAGGGAATTTGGGGTTGAGGAAGCACAGCGCTGGCAGGAATAAGATGAGT 2640
Db 2581 ACTCATGTCAGGGAATTTGGGGTTGAGGAAGCACAGCGCTGGCAGGAATAAGATGAGT 2640
Qy 2641 GAGACAGCAAGGCTATTGGAATCCACACCCAGAACCAAGGGGTGAGCCCTGGACACC 2700
Db 2641 GAGACAGCAAGGCTATTGGAATCCACACCCAGAACCAAGGGGTGAGCCCTGGACACC 2700
Qy 2701 TCACCCAGGATGGCTTCTTTTCACTCTCTTTCAGATCTGGGCGAGGTGAGGACCT 2760
Db 2701 TCACCCAGGATGGCTTCTTTTCACTCTCTTTCAGATCTGGGCGAGGTGAGGACCT 2760
Qy 2761 CATTTCTCAGAGGTTGACTCAGGTCACTAGGACCCCTCTGTTCTAAGACAGAGCG 2820
Db 2761 CATTTCTCAGAGGTTGACTCAGGTCACTAGGACCCCTCTGTTCTAAGACAGAGCG 2820
Qy 2821 GTCCAGGATCTGCCATGCGTTCGGGTGAGAAATCAGGAGGACTGAGGGTACCCAG 2880
Db 2821 GTCCAGGATCTGCCATGCGTTCGGGTGAGAAATCAGGAGGACTGAGGGTACCCAG 2880
Qy 2881 GACCAGAACTGAGGAGGACTGCACAGAAATCAGCCCTGCTCTGTCACCCAGAG 2940
Db 2881 GACCAGAACTGAGGAGGACTGCACAGAAATCAGCCCTGCTCTGTCACCCAGAG 2940

Db	5101	TAAGACTCTTGGTTTTATTTCAGATTGGAAATCCATTCTATTTTTGTGAATTGGGATAAT	5160
Qy	5161	ACAGCAGTGGAAATAGTACTTTAGAAATGCTGAAAATCAGCAGCTAAATAGATGAGATAA	5220
Db	5161	ACAGCAGTGGAAATAGTACTTTAGAAATGCTGAAAATCAGCAGCTAAATAGATGAGATAA	5220
Qy	5221	AGAACTAAAGAAATTAAGAGATAGTCAATTCCTGGCTTATACCTCAGTCTATTCTGTAAA	5280
Db	5221	AGAACTAAAGAAATTAAGAGATAGTCAATTCCTGGCTTATACCTCAGTCTATTCTGTAAA	5280
Qy	5281	ATTTTAAAGATATATGCATACCTGGATTTCTTGGCTCTTTTGAGAAATGTAAGAGAAAT	5340
Db	5281	ATTTTAAAGATATATGCATACCTGGATTTCTTGGCTCTTTTGAGAAATGTAAGAGAAAT	5340
Qy	5341	TAAATCTGAATAAAGAAATTCCTCTGTTTCACCTGGCTCTTTTCTCTCCATGCACTGAGCA	5400
Db	5341	TAAATCTGAATAAAGAAATTCCTCTGTTTCACCTGGCTCTTTTCTCTCCATGCACTGAGCA	5400
Qy	5401	TCGTCTTTTGGAAAGCCCTGGGTTAGTAGTGGAGATGCTTAAGGTAAGCCAGACTCATC	5460
Db	5401	TCGTCTTTTGGAAAGCCCTGGGTTAGTAGTGGAGATGCTTAAGGTAAGCCAGACTCATC	5460
Qy	5461	CCACCATAAGGTCGTAGAGTCTAGGAGCTGCAGTCACTAATCGAGGTGCCAAGATGTC	5520
Db	5461	CCACCATAAGGTCGTAGAGTCTAGGAGCTGCAGTCACTAATCGAGGTGCCAAGATGTC	5520
Qy	5521	CTCTAAAGATGTAGGAAAAGTGCAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG	5580
Db	5521	CTCTAAAGATGTAGGAAAAGTGCAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG	5580
Qy	5581	GTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTTGGGAAACTGCAGTTCTCTCT	5640
Db	5581	GTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTTGGGAAACTGCAGTTCTCTCT	5640
Qy	5641	GGGGAGCTGATTG 5654	
Db	5641	GGGGAGCTGGCTG 5654	

4

RESULT

Q32351 Q32351 standard; DNA; 2419 BP.

ID Q32351 AC Q32351

DT 22-APR-1993 (first entry)

DE Antigen E gene.

EW Stable; antigen; E; D; F; A; human; melanoma; cell line; M22-MEL;

KW cytolytic T cell; MEL3.1; open reading frame; homology; MAGE;

KW melanoma antigen; ss.

OS Homo sapiens.

PN WO9220356-A.

PD 26-NOV-1992.

PF 22-MAY-1992; U04354.

PR 23-MAY-1991; US-705702.

PR 09-JUL-1991; US-728838.

PR 23-SEP-1991; US-764364.

PR 12-DEC-1991; US-807043.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;

PI Van den Eynde B, Van der Bruggen P, Van Pel A;

PT WPI; 92-413460/50.

PT Nucleic acid mol. encoding a human tumour rejection antigen

PT precursor - useful as an immunostimulant in a vaccine for

PT treating and preventing cancers, also useful in diagnosis .

PS Disclosure: Page 69-70; 14pp; English.

CC This sequence encodes the stable antigen E. This antigen is expressed

CC along with antigens "D, F and A" by the human melanoma cell line M22-

CC MEL. These antigens are all recognised by cytolytic T cells. A

CC subline of M22-MEL is MEL3.1 which only expresses antigen E. This

CC cell line was chosen as a source for the isolation of this sequence.

CC This sequence was found to contain three exons. The open reading frame

CC for antigen E was contained within the first two exons. An ATG is

CC located at position 66 of exon 3 and is followed by an 828 base pair

CC 'reading frame. The three exons contain 65, 73 and 1551 base pairs.

QY 4156 TATCTCGAGTCTTGTTCCTCGAGCAGTAACTACTAAGAAAGTGGCTGATTGGTTGGTTT 4215
DB 901 TATCTCGAGTCTTGTTCCTCGAGCAGTAACTACTAAGAAAGTGGCTGATTGGTTGGTTT 960
QY 4216 TCTGCTCTCAATATCGAGCAGGAGCCAGTCAAAAGGAGCAAAATCTCGAGAGTCT 4275
DB 961 TCTGCTCTCAATATCGAGCAGGAGCCAGTCAAAAGGAGCAAAATCTCGAGAGTCT 1020
QY 4276 CATCAAAAATACAGCACTGTTTTCTCTGAGATCTTCGGCAAGCCCTGAGTCCCTTGA 4335
DB 1021 CATCAAAAATACAGCACTGTTTTCTCTGAGATCTTCGGCAAGCCCTGAGTCCCTTGA 1080
QY 4336 GCTGGTCTTGGCATGAGTGAAGAAAGCAGACCCACCCGCGCACTCTATGTCCTTGT 4395
DB 1081 GCTGGTCTTGGCATGAGTGAAGAAAGCAGACCCACCCGCGCACTCTATGTCCTTGT 1140
QY 4396 CACCTGCCTAGTCTCTCTATGATGGCTGCTGGGTGATAATCAGATCATGCCCAAGAC 4455
DB 1141 CACCTGCCTAGTCTCTCTATGATGGCTGCTGGGTGATAATCAGATCATGCCCAAGAC 1200
QY 4456 AGGCTTCTGATAATGTCCTGTCATGATGCAATGAGGCGGCCATGCTCCTCGAGA 4515
DB 1201 AGGCTTCTGATAATGTCCTGTCATGATGCAATGAGGCGGCCATGCTCCTCGAGA 1260
QY 4516 GGAATCTGGGAGGAGTGAAGTGTATGAGTGTATGATGGAGGAGCAGACGCTCA 4575
DB 1261 GGAATCTGGGAGGAGTGAAGTGTATGAGTGTATGATGGAGGAGCAGACGCTCA 1320
QY 4576 TGGGAGCCAGGAGTCTCTACCAACAGATTGTTGTCAGGAAAGTACTCTGGAGTACGG 4635
DB 1321 TGGGAGCCAGGAGTCTCTACCAACAGATTGTTGTCAGGAAAGTACTCTGGAGTACGG 1380
QY 4636 CAGGTGCCGAGCAGTATGCCGACGCTATGATGCTCTGGGGTCCAAAGGCGCTCGCT 4695
DB 1381 CAGGTGCCGAGCAGTATGCCGACGCTATGATGCTCTGGGGTCCAAAGGCGCTCGCT 1440
QY 4696 GAAACAGCTATGTGAAGTCTTGAAGTGTATGATCAAGGTGAGTCAAGAGTTCGCTTT 4755
DB 1441 GAAACAGCTATGTGAAGTCTTGAAGTGTATGATCAAGGTGAGTCAAGAGTTCGCTTT 1500
QY 4756 TTCTTCCCATCTCGCTGAACAGCTTTGAGAGGAGGAGGAGGAGTCTGAGCATGA 4815
DB 1501 TTCTTCCCATCTCGCTGAACAGCTTTGAGAGGAGGAGGAGGAGTCTGAGCATGA 1560
QY 4816 GTTGACGCAAGCCAGTGGGAGGGGAGTGGCCAGTGCACCTTCCAGGGCGCGTCCA 4875
DB 1561 GTTGACGCAAGCCAGTGGGAGGGGAGTGGCCAGTGCACCTTCCAGGGCGCGTCCA 1620
QY 4876 GCAGCTTCCCTGCTGCTGACATGAGGCCCATTTCTCACTCTGAAGAGAGCGGTGAG 4935
DB 1621 GCAGCTTCCCTGCTGCTGACATGAGGCCCATTTCTCACTCTGAAGAGAGCGGTGAG 1680
QY 4936 TGTCTCAGTAGTGGTTCTGTTCTTATGAGTGGTGAATTTATCTTCTTCTCTT 4995
DB 1681 TGTCTCAGTAGTGGTTCTGTTCTTATGAGTGGTGAATTTATCTTCTTCTCTT 1740
QY 4996 TTGGAATCTTCAAAATGTTTTTTTTTAAGGATGTTGTAATGAATTCAGCATCCAGTT 5055
DB 1741 TTGGAATCTTCAAAATGTTTTTTTTTAAGGATGTTGTAATGAATTCAGCATCCAGTT 1800
QY 5056 TATGAATGACAGCAGTCAACAGTCTGCTGTATATATGTTTAAAGGTAAGAGTCTTGT 5115
DB 1801 TATGAATGACAGCAGTCAACAGTCTGCTGTATATATGTTTAAAGGTAAGAGTCTTGT 1860
QY 5116 TTATTCAGATGGGAATCCATCTATTTTGTGAATTTGGGATAATTAACAGCAGTGAATA 5175
DB 1861 TTATTCAGATGGGAATCCATCTATTTTGTGAATTTGGGATAATTAACAGCAGTGAATA 1920
QY 5176 AGTACTTGAATAATGAAATGAGCAGTAAATAGATGAGATAAGAACTAAAGAAAT 5235
DB 1921 AGTACTTGAATAATGAAATGAGCAGTAAATAGATGAGATAAGAACTAAAGAAAT 1980
QY 5236 AAGAGATAGTCAATCTTCCCTTATACCTCAGTCTATTCTGTAAAAATTTTTAAAGATATA 5295

DB 1981 AAGAGATAGTCAATCTTCCCTTATACCTCAGTCTATTGTAAAAATTTTTAAAGATATA 2040
QY 5296 TGCATACCTGGATTTCTTGGCTTCTTTCAGAAATGTAAGAGAAATTAATCTCAATAAG 5355
DB 2041 TGCATACCTGGATTTCTTGGCTTCTTTCAGAAATGTAAGAGAAATTAATCTCAATAAG 2100
QY 5356 AATTCTTCTGTTACTGCTCTTTTCTTCTCCATGCACTGAGCATCTCTTTTTCGAAG 5415
DB 2101 AATTCTTCTGTTACTGCTCTTTTCTTCTCCATGCACTGAGCATCTCTTTTTCGAAG 2160
QY 5416 GCCCTGGGTAGTGGAGATGCTAAGTAAGCCAGACATCATACCCCATAGGTCG 5475
DB 2161 GCCCTGGGTAGTGGAGATGCTAAGTAAGCCAGACATCATACCCCATAGGTCG 2220
QY 5476 TAGAGTCTAGGAGTGCAGTCACTAAGTGAAGTGGCAAGATGCTCTCTAAAGATGAG 5535
DB 2221 TAGAGTCTAGGAGTGCAGTCACTAAGTGAAGTGGCAAGATGCTCTCTAAAGATGAG 2280
QY 5536 GAAAAGTGAAGAGGGGTGAGGTGTTGGGCTTTCGGAACCTGCAAGTGTGAGTGTCAATGC 5595
DB 2281 GAAAAGTGAAGAGGGGTGAGGTGTTGGGCTTTCGGAACCTGCAAGTGTGAGTGTCAATGC 2340
QY 5596 CTTGAGCTGGGCAATTTGGGCTTTCGGAACCTGCAAGTGTGAGTGTGAGTGTGAGTGT 5655
DB 2341 CTTGAGCTGGGCAATTTGGGCTTTCGGAACCTGCAAGTGTGAGTGTGAGTGTGAGTGT 2400
QY 5656 AATGATCTTGGGTGGATCC 5674
DB 2401 AATGATCTTGGGTGGATCC 2419

RESULT 5
Q72476 standard; DNA; 2419 BP.
ID Q72476;
AC Q72476;
DT 21-JUN-1995 (first entry)
DE Tumour rejection antigen E encoding DNA.
KW Tumour rejection antigen E; melanoma antigen-3; MAGE-3;
KW cancer; cytolytic T cells; antigen D; human leucocyte antigen; ss.
OS Homo sapiens.
PN W09423031-A.
PF 17-MAR-1994; U02877.
PR 26-MAR-1993; US-037230.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon-falleur T, Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;
DR WPI; 94-333192/41.
PT New tumour rejection antigen precursor MAGE3 - useful in
PT treatment and diagnosis of cancer
PS Disclosure; Page 58; 105pp; English.
CC Q72476 encodes tumour rejection antigen E, another sequence
CC Q72470 encodes melanoma antigen-3 (MAGE-3) a tumour rejection
CC antigen precursor. Melanomas characterised by the expression of
CC MAGE-3 can be detected, or monitored, by contacting a test sample
CC with an agent that can recognise MAGE-3. The melanoma can be treated
CC by the administration of cytolytic T cells specific for the complex of
CC antigen D (the mature rejection antigen derived from MAGE-3) and a human
CC leucocyte antigen (esp. HLA-A1).
SQ Sequence 2419 BP; 562 A; 581 C; 677 G; 599 T;

Query Match 42.6%; Score 2419; DB 1; Length 2419;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3256 GGATCCAGGCTCCGCCAGGAAAAATAAGGGCCCTCGTGCAGACAGAGGGGTCTATCC 3315
DB 1 GGATCCAGGCTCCGCCAGGAAAAATAAGGGCCCTCGTGCAGACAGAGGGGTCTATCC 60
QY 3316 ACTGCATGAGTGGGGATGTACAGAGTCCAGGCCACCCCTCTCTGTAGACTGAGAAGC 3375
DB 61 ACTGCATGAGTGGGGATGTACAGAGTCCAGGCCACCCCTCTCTGTAGACTGAGAAGC 120

CC MAGE-1 antigens. Compsns. containing these peptides can be
CC administered, as a vaccine to patients susceptible to MAGE
CC associated tumours, e.g. melanomas.
SQ Sequence 2420 BP; 562 A; 582 C; 677 G; 599 T;

Query Match 42.4%; Score 2408; DB 1; Length 2420;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2419; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 3256 GGATCCAGGCCCTGCCAGGAAAAATATTAAGGCCCTGGCTGAGACAGAGGGGTCTATCC 3315
D 1 GGATCCAGGCCCTGCCAGGAAAAATATTAAGGCCCTGGCTGAGACAGAGGGGTCTATCC 60
QY 3316 ACTGCATCAGAGTGGGGATGTACAGAGTCCAGGCCACCCCTCCTGGTAGCAGTGAAGC 3375
D 61 ACTGCATCAGAGTGGGGATGTACAGAGTCCAGGCCACCCCTCCTGGTAGCAGTGAAGC 120
QY 3376 CAGGCTGTGTTGGGCTGCACCCCTGAGGCCCGGTGGATTCCTTCCTGGAGCTCCA 3435
D 121 CAGGCTGTGTTGGGCTGCACCCCTGAGGCCCGGTGGATTCCTTCCTGGAGCTCCA 180
QY 3436 GGAACCCAGGCTGAGGCCCTGGCTGAGACAGTATCCTCAGGTCACAGACAGAGGATG 3495
D 181 GGAACCCAGGCTGAGGCCCTGGCTGAGACAGTATCCTCAGGTCACAGACAGAGGATG 240
QY 3496 CACAGGGTGTGCCAGCAGTGAATGTTGCCCTGAATGCACACCAAGGCCCCACCTGCCA 3555
D 241 CACAGGGTGTGCCAGCAGTGAATGTTGCCCTGAATGCACACCAAGGCCCCACCTGCCA 300
QY 3556 CAGGACATAGGACTCCACAGAGTCTGGCTCACCCTCCTACTGTCTAGTCTCTAGTAA 3615
D 301 CAGGACATAGGACTCCACAGAGTCTGGCTCACCCTCCTACTGTCTAGTCTCTAGTAA 360
QY 3616 CGACCTCTGCTGGCGGCTGACCTGAGTACCTCTCCTCCTCCTCCTCCTCCTCCTCCT 3675
D 361 CGACCTCTGCTGGCGGCTGACCTGAGTACCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 420
QY 3676 GGGACAGCCCAACCCAGAGGACAGGATTCCTGGAGGCCACAGAGGACACCAAGGAA 3735
D 421 GGGACAGCCCAACCCAGAGGACAGGATTCCTGGAGGCCACAGAGGACACCAAGGAA 480
QY 3736 GATCTGTAAGTAGGCTTTGTTAGAGTCTCCAAAGTTTCAGTTCAGTTCAGTTCAGTTC 3795
D 481 GATCTGTAAGTAGGCTTTGTTAGAGTCTCCAAAGTTTCAGTTCAGTTCAGTTCAGTTC 540
QY 3796 CACACTCCCTCTCTCCAGGCCCTGGGCTCTCATTTGCCAGCTCCTGCCACACTCCT 3855
D 541 CACACTCCCTCTCTCCAGGCCCTGGGCTCTCATTTGCCAGCTCCTGCCACACTCCT 600
QY 3856 GCCTGCTGCCCTGACGAGATCATCTCTCTGAGCAGAGGAGTCTGCACCTGCAAGCC 3915
D 601 GCCTGCTGCCCTGACGAGATCATCTCTCTGAGCAGAGGAGTCTGCACCTGCAAGCC 660
QY 3916 TGAGGAAGCCCTTGAGGCCCAACAGAGGCCCTGGGCTGTGTGTGTCAGGCTGCCAC 3975
D 661 TGAGGAAGCCCTTGAGGCCCAACAGAGGCCCTGGGCTGTGTGTGTCAGGCTGCCAC 720
QY 3976 CTCCTCTCTCTCTCTCTGCTGGGACCTGGAGAGCTGCCACTGCTGGGTCAC 4035
D 721 CTCCTCTCTCTCTCTCTGCTGGGACCTGGAGAGCTGCCACTGCTGGGTCAC 780
QY 4036 AGATCCTCCCTCAGAGTCTCAGGAGGCTCCGCCCTTTCCCACTACCATCACTTCCTCG 4095
D 781 AGATCCTCCCTCAGAGTCTCAGGAGGCTCCGCCCTTTCCCACTACCATCACTTCCTCG 840
QY 4096 ACAGAGCAACCCAGTAGGCTTCAGCAGCCGTCAGAGAGGGGCCAAGCACTCTTG 4155
D 841 ACAGAGCAACCCAGTAGGCTTCAGCAGCCGTCAGAGAGGGGCCAAGCACTCTTG 900
QY 4156 TATCTGGAGTCTCTGTTCCAGCAGTAACTCACTAAGAGCTGCTGATTTGGTGT 4215
D 901 TATCTGGAGTCTCTGTTCCAGCAGTAACTCACTAAGAGGCTGCTGATTTGGTGT 960
```

```
QY 4216 TCTCTCTCAAAATATCAGCCAGGAGCCAGTCACAAAGCAGAAATGCTGGAGAGTGT 4275
D 961 TCTCTCTCAAAATATCAGCCAGGAGCCAGTCACAAAGCAGAAATGCTGGAGAGTGT 1020
QY 4276 CATCAAAATTTACAAGCACTGTTTCTCTGAGATCTTCGGCAAGCCCTCTGAGTCTTGCA 4335
D 1021 CATCAAAATTTACAAGCACTGTTTCTCTGAGATCTTCGGCAAGCCCTCTGAGTCTTGCA 1080
QY 4336 GCTCGTCTTTGGCATTCACGTAAGGAGCAGACCCACCGCCACCTCTATGCTCTTGT 4395
D 1081 GCTCGTCTTTGGCATTCACGTAAGGAGCAGACCCACCGCCACCTCTATGCTCTTGT 1140
QY 4396 CACCTGCTAGTCTCTCTATGATGCTGCTGGTGATTAATCAGATCATGCCCAAGC 4455
D 1141 CACCTGCTAGTCTCTCTATGATGCTGCTGGTGATTAATCAGATCATGCCCAAGC 1200
QY 4456 AGGCTTCTGATTAATGCTGCTGATGATGAGGCGGCCATGCTCTCTGAGGA 4515
D 1201 AGGCTTCTGATTAATGCTGCTGATGATGAGGCGGCCATGCTCTCTGAGGA 1260
QY 4516 GGAATCTGGAGCAGCTGATGATGAGGCTGATGATGAGGCGGCCATGCTCTGAGGA 4575
D 1261 GGAATCTGGAGCAGCTGATGATGAGGCTGATGATGAGGCGGCCATGCTCTGAGGA 1320
QY 4576 TGGGAGCCAGGAGCTGCTCACCAGATTTGGTCAGGAAAGTACCTGGAGTA-CG 4634
D 1321 TGGGAGCCAGGAGCTGCTCACCAGATTTGGTCAGGAAAGTACCTGGAGTACCG 1380
QY 4635 GCAGGTGCCGAGCAGTGAATCCCGCACGCTATGATTCCTGTGGGTCGCAAGGCCCTGC 4694
D 1381 GCAGGTGCCGAGCAGTGAATCCCGCACGCTATGATTCCTGTGGGTCGCAAGGCCCTGC 1440
QY 4695 TGAACACAGCTATGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 4754
D 1441 TGAACACAGCTATGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1500
QY 4755 TTTCTTCCCATCCCTGGCTGAAGCAGCTTTGAGAGGAGGAGGAGGAGTCTGAGCATG 4814
D 1501 TTTCTTCCCATCCCTGGCTGAAGCAGCTTTGAGAGGAGGAGGAGGAGTCTGAGCATG 1560
QY 4815 AGTTGACGCAAGGCCAGTGGAGGGGAGTGGGCCAGTTCACCTTCAGGGCCGCTGCC 4874
D 1561 AGTTGACGCAAGGCCAGTGGAGGGGAGTGGGCCAGTTCACCTTCAGGGCCGCTGCC 1620
QY 4875 AGCAGCTTCCCTCCCTGCTGATGATGAGGCCCATCTTTCACCTCTGAAGAGAGGGTCA 4934
D 1621 AGCAGCTTCCCTCCCTGCTGATGATGAGGCCCATCTTTCACCTCTGAAGAGAGGGTCA 1680
QY 4935 GTGTTCTCAGTAGTGTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCT 4994
D 1681 GTGTTCTCAGTAGTGTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCT 1740
QY 4995 TTTGGAATTTGTTCAAAATGTTTTTTTAAAGGATGTTGATGAATGACATTCAGCATCCAA 5054
D 1741 TTTGGAATTTGTTCAAAATGTTTTTTTAAAGGATGTTGATGAATGACATTCAGCATCCAA 1800
QY 5055 TTATGAATGACAGCAGTACACACTTCTGCTATATAGTTTAAAGGTAAGAGTCTTGTGT 5114
D 1801 TTATGAATGACAGCAGTACACACTTCTGCTATATAGTTTAAAGGTAAGAGTCTTGTGT 1860
QY 5115 TTTATTTCAGATTGGGAATCCATTCTATTTTGTGAATTTGGATTAATACACAGCTGGAAT 5174
D 1861 TTTATTTCAGATTGGGAATCCATTCTATTTTGTGAATTTGGATTAATACACAGCTGGAAT 1920
QY 5175 AAGTACTTAGAAATGTTGAAATGAGCAGTAAATAGATGATGATGATGATGATGATGATG 5234
D 1921 AAGTACTTAGAAATGTTGAAATGAGCAGTAAATAGATGATGATGATGATGATGATGATG 1980
QY 5235 TAACAGATAGTCAATCTTGGCTTATACCTCAGTCTATCTGTAAGATTTTTTAAAGATAT 5294
D 1981 TAACAGATAGTCAATCTTGGCTTATACCTCAGTCTATCTGTAAGATTTTTTAAAGATAT 2040
```

QY 5295 ATGCATACCTGGATTCTTCTGGCTTCTTTGAGAAATGTAAGAGAAATTAATCTGAATAAA 5354
 |||||
 Db 2041 ATGCATACCTGGATTCTTCTGGCTTCTTTGAGAAATGTAAGAGAAATTAATCTGAATAAA 2100
 |||||
 QY 5355 GAATTCCTCTCTTCTACCTGGCTTCTTCTCTCCATGCACTGAGCATCTGCTTTTGGAA 5414
 |||||
 Db 2101 GAATTCCTCTCTTCTACCTGGCTTCTTCTCTCCATGCACTGAGCATCTGCTTTTGGAA 2160
 |||||
 QY 5415 GGCCTGGGTAGTAGTGGAGATGCTAAGTAAGCCAGACTCATACCCACCCATAGGCTC 5474
 |||||
 Db 2161 GGCCTGGGTAGTAGTGGAGATGCTAAGTAAGCCAGACTCATACCCACCCATAGGCTC 2220
 |||||
 QY 5475 GTAGAGCTAGAGCTGAGTACGTAATCGAGGTGGCAAGATGCTCTTAAGATGTAG 5534
 |||||
 Db 2221 GTAGAGCTAGAGCTGAGTACGTAATCGAGGTGGCAAGATGCTCTTAAGATGTAG 2280
 |||||
 QY 5535 GGAAGAAGTGAGAGAGGGGTGAGGGTGGGGCTCGGGGTGAGAGTGGTGGAGTGCAATG 5594
 |||||
 Db 2281 GGAAGAAGTGAGAGAGGGGTGAGGGTGGGGCTCGGGGTGAGAGTGGTGGAGTGCAATG 2340
 |||||
 QY 5595 CCCTGAGCTGGGGCATTTTGGGCTTTGGGAACTGCACTTCTTCTGGGGGAGCTGATTG 5654
 |||||
 Db 2341 CCCTGAGCTGGGGCATTTTGGGCTTTGGGAACTGCACTTCTTCTGGGGGAGCTGATTG 2400
 |||||
 QY 5655 TAATGATCTGGGTGGATCC 5674
 |||||
 Db 2401 TAATGATCTGGGTGGATCC 2420
 |||||

RESULT 9

Q72478
 ID Q72478 standard; DNA: 4157 BP.
 AC Q72478;
 DT 22-JUN-1995 (first entry)
 DE Tumour rejection antigen MAGE-2 gene.
 KW Tumour rejection antigen; melanoma antigen-2; MAGE-2; MAGE-3;
 KW cancer; cytolytic T cells; antigen D; human leucocyte antigen;
 KW ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 Key 2598..3542
 cds /*tag= a
 FT
 FT
 PN W09423031-1-A.
 PD 13-OCT-1994.
 PF 17-MAR-1994; U02877.
 PR 26-MAR-1993; US-037230.
 PA (LUDW-) LUDWIG INST CANCER RES.
 PI Boon-faller T, Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;
 DR WPI: 94-333192/41.
 PT New tumour rejection antigen precursor MAGE3 - useful in
 treatment and diagnosis of cancer
 PS Example 32; Page 61; 105pp; English.
 CC Q72478 is the gene which contains the coding sequence for melanoma
 antigen-2 (MAGE-2). Another melanoma antigen precursor. Melanomas
 Q72470, this is a tumour rejection antigen precursor. Melanomas
 characterised by the expression of MAGE-3 can be detected, or
 monitored, by contacting a test sample with an agent that can
 recognise MAGE-3. The melanoma can be treated by the administration
 of cytolytic T cells specific for the complex of antigen D (the
 mature rejection antigen derived from MAGE-3) and a human leucocyte
 antigen (esp. HLA-A1).
 SQ Sequence 4157 BP; 953 A; 1134 C; 1185 G; 885 T;

Query Match 36.3%; Score 2058.8; DB 1; Length 4157;
 Best Local Similarity 75.3%; Pred. No. 0;
 Matches 3172; Conservative 0; Mismatches 837; Indels 201; Gaps 41;

QY 1351 CCATCGCCTCCCATCTTGGCAAGATCCGGT-TGCCCTCTCTCAACCCAGGGAAG 1409
 |||||
 Db 1 CCATCCAGATCCCATCCGGCGAGNATCCGGTTCACCCCTTCCCGTGAACCCAGGGAAG 60
 |||||
 QY 1410 CCCTGGTAGGCCCGATGTGAACACCACTGACTTGAACCTCACAGATCTGAGAGAAGCCAGG 1469
 |||||

Db 61 TCACGG--GCCCGGATGTGACGCCACTGACTTGCACATTGGAGGTGAGAGGACAGCGAGA 118
 |||||
 QY 1470 TTCAATTAATGTTCTGAGGGGCGCTTGAGATCCACTGAGGGAGTGTGTTTAGGCTCT 1529
 |||||
 Db 119 TTC-----TCGCCCTGAGCAACGGCTGACGTGCGGGAGGGAAGCAGCGCAGGCTCC 172
 |||||
 QY 1530 GTGAGAGGCAAGGTGAGATGCTGAGGGAGGACTGAGGAGGCACACACCCAGGTAGATG 1589
 |||||
 Db 173 GTGAGAGGCAAGGTGAGAGCGCGAGGAGGACTGAGGGGGCTCACCCAGACAGAGG 232
 |||||
 QY 1590 GCCCAAAATGATCAGTACCACTGCTGCTCCAGCCCTGGACCACCCCGCCAGGACAGA 1649
 |||||
 Db 233 GCCCCCAATTAATCAACGGCTGCTGCTGCCGGCCCTGGACCACCCCTGCGAGGGGAAGA 292
 |||||
 QY 1650 TGTCTCAG:-----CTGGAACACCCCGCTCCCGTCCCTGCTCCACTTGAACCCACAGG 1702
 |||||
 Db 293 CTTCTCAGGCTCAGTCCGCAACCACTCACCCGCCACCCCGCCGCTTAAACCGCAGGG 352
 |||||
 QY 1703 CAATCTGTAGTCAATAGCTTA-TGTGACCGGGGCGAGGTTGGTCAGGAGAGGAGGCCCA 1761
 |||||
 Db 353 AACTCTGGCGTAAGAGCTTGTGTACCAAGGCGAGGCGTGGTTAGAAAGTGC----- 404
 |||||
 QY 1762 GGCATCAAGGTCCAGATCCCGCCCGGCAATTAGGTGTCAGGACCCCTGGGAGGGAAGT 1821
 |||||
 Db 404 -----TCAGGGCCGAGACTCAGCCAGGAATCAAGGTCAAGGCCCAAGAGGGGACTGAGG 459
 |||||
 QY 1822 TTCCCCACCCACACCTGCTCTCATCTCC----- 1852
 |||||
 Db 460 CAACCCACCCCTACCCCTACCTACCAATCCCATCCCAACCAACCAACCCCAACCCATCC 519
 |||||
 QY 1852 -----ACGCCCACCCCACTACATTCCTACCTACCCCTACCCCAACCACTCTTGT 1906
 |||||
 Db 520 CTCAAAACCAACCCCAACCCCACTTCCCATCTCTCTCCCAACCACTCTCTGG 579
 |||||
 QY 1907 CAGAAAT-----CCCTGCTGTCAACCCAGGAGCCAGGGGATGGCGCCAGGCA 1956
 |||||
 Db 580 CAGAAATCCGGCTTTGCCCTGCAATCAACCCAGGAAGTCCGGGAATGGCGCCCAAGCA 639
 |||||
 QY 1957 CTCGGATCTTGACCTCCCATCCAGGCTGTGATGAGGAAGGG-----CTTGAAC 2008
 |||||
 Db 640 CGCGGATCTGACGTTCACATGACG-CTAAGGAGGGAAGGGTTGGGTCTCTGTGAGT 698
 |||||
 QY 2009 AGGGCTCAGGGAGCAGAGGAGGCG-----CCTACTCGAGATGAGGAGGCGCTCAGA 2063
 |||||
 Db 699 ATGGCTTTGGGATGCAAGAGGAAGGCGCCAGGCTCTCTGGAAGACAGTGGAGTCTTAGG 758
 |||||
 QY 2064 GGACCCAGCACCCTAGGA-----CACCGCACCCTGTCTGAGACTGAGGCTGCCAC 2114
 |||||
 Db 759 GGACCCAGCATCCAGGACAGGCGGCCCACTGTACCCCTGTCTCAAACTGAGCCACCTT 818
 |||||
 QY 2115 TTCTGSCCTCAAGAATCAGAAAGATGGGACTCAGATTGCGGGGTGGGACCCAGGCG 2174
 |||||
 Db 819 TCATTACGCCGAGGGAATCCTAGGATGAGACCCACTTCAGGGGTTGGGCGCCAGCCT 878
 |||||
 QY 2175 TCAAGGCTTACCGGAGGAAGAGGAGGACTCAGGGGACCTTTGGAATCCAGATCAG 2234
 |||||
 Db 879 GCGAGGAGTCAAGGGGAGGAAGAGGAGGACTGAGGGGACCTTTGGAGTCCAGATCAG 938
 |||||
 QY 2235 TGTGAGACTCGGCCCTGAGAGGTCCAGGCGAGGTGGCCACATATGCCCCATATTTCTTG 2294
 |||||
 Db 939 TGGCAACCTTGG-GCTGGGGGATCTTGGGCACAGTGGCGGAATGTGCCCGCTGCTCATTTG 997
 |||||
 QY 2295 CATCTTTGAGGTGAC----AGCACAGAGCTGTGCTGAGAAGTGGGGCTCAGGTCAC 2350
 |||||
 Db 998 CACCTTCAGGGTGACAGAGAGTTGAGGGCTGTGGTCTGAGGGCTGGGACTTCAGGTGAGC 1057
 |||||
 QY 2351 AGAGGAGGAGTTCAGGATCCATATGGCCCAAGATGTGCCCTTCATGAGGAGTGGG 2410
 |||||
 Db 1058 AGAGGAGGAATCCAGGATCTGCCGAGCCCAAGGTGTGCCCTTCATGAGGACTCCCC 1117
 |||||
 QY 2411 ATATCCCCGGCTCAGAAAGAAGGAGCTCCACACAGTCTGGCTGTCCCCCTTTTAGTAGCTC 2470
 |||||

Db 1118 ATACCCCGGCCCCAGAAAGAGGATGCCACAGAGTCTGGAAGTAAATGTCTTAGCTC 1177
Qy 2471 TAGGGACACAGATCAGGAGTGGCGGTATGTCTCATCTACTTGTACACAGGACAGGAA 2530
Db 1178 TGGGGAACCTGATCAGGATGGCCCTAAGTGACAATCTCATTTGTACACAGGACAGG 1237
Qy 2331 GTTGGGGGCGCTCAGGAGATGGGCTCTTGGGGTAAAGGGGGATGTCTACTCATGTCA 2590
Db 1238 GTTGGGGAACCTCAGGAGATAAGGTGTGGTGTAAAGAGGAGCTGTCTGCTCATTTCA 1297
Qy 2591 GGGAAATGGGGTTGAGGAACACAGGCGCTGGCAGGATAAAGATGAGTGAGACACACA 2650
Db 1298 GGGGGTTCCCGCTTGGAGAAAGGAGGAGTCCCTGGCAGGAGTAAAGATGAGTAACCCACAGG 1357
Qy 2651 AGGCTATTGGAAATCCACACCCCAACAAAGGGGTGAGCCCTGGACACCTCACCCACAGGA 2710
Db 1358 AGGCATCATACCTTACCTTGAGAACCAAGGGGTGAGCCCTGGACACACCCAGCTGGG 1417
Qy 2711 T-----GTGGCTTCTTTTCACTCTCTGTTCAGATCTGGGCGAGGTGAGGACCTCAT 2763
Db 1418 TAACAGGATGTGGCCCTCTCTCACTTGTCTTCCAGATCTCAGGGAGTTGATGACCTTGT 1477
Qy 2764 TCTCAGAGGTGACTCAGGTCAAGTAGGACCCCACTCTGGTCTAAGACAGAGCGGTC 2823
Db 1478 TTTCAGAAAGTGACTCAGTCAACACAGGG---CCCTCTGGTGCAGAGATGCGAGTGGT 1534
Qy 2824 CCAGGATCTGCCATCGCTTGGGTGAGGAACATCAGGAGGAGCTGAGGGTACCCACAGGAC 2883
Db 1535 CTAGGATCTGCCAAGCATCCAGGTGGAGGCTCAGGTAGGATGAGGGTACCCCTGGG 1594
Qy 2884 CAGAACTG-AGGGAGACTGCACAGAAATCAGCCCTGCCCTCTGTACACCCACAGAG 2942
Db 1595 CAGAAATCAGCAAGGGGCCCCATAGAAATCTGCCCTGCCCTCGGCTTACTTCAGAGAC 1654
Qy 2943 CATGGCTGGCCCTCTGCCAGTCTCTCGGTATCTCTGGATCATGTGTCAGGAC 3002
Db 1655 CTTGGCAGGGCTCTCAGCTGAAGTCCCTCATTTAT-CTGGGATCTTTGATGTCAAGGAA 1713
Qy 3003 GGGAGGCTTGGTCTGAGAAGGCTGGGCTCAGGTCAAGTGAAGGAGCGTCCACAGGCCCT 3062
Db 1714 GGGAGGCTTGGTCTGAAGGGCTGGATCAGGTCAAGTGAAGGAGGGTCTCAGGCCCT 1773
Qy 3063 GCCAGGATCAAGGTGAGGACCAAGCGGACCTCACCCAGGACACATTAATTCATGA 3122
Db 1774 GCCAGGATGAGCTGAGGACCAAGCGGACTCTGTCACCCAGGACACCTGGACTCCAATGA 1833
Qy 3123 ATTTGATATCTCTGTCTGCCCTTCCCAAGGACCTAGGACGCTGTCGCGAGATGTTGT 3182
Db 1834 A-TTGGATCTCTCTGTCTTCCGCGGAGGACCTGGTCAAGTATGGCCAGATGTGGT 1892
Qy 3183 CCCCTC-CTGTCTCTTCCATCTTATCATGATGTGAACCTTTC-ATTGGATTTCTCA 3239
Db 1893 CCCCTCTCTCTCTGTACCATATCAGGATGTGAGTCTTGCATGAGAGATTTCTCA 1952
Qy 3240 GACCAAAAGGCGAGATCCAGGCCCTTCCAGGAAATAATTAAGGGCCCTGGGTGAGA 3299
Db 1953 AGCCAGAAAGGGTGGAT-TAGGCCCTACAAGGAGAAAGGTGAGGGCCCTGAGTGAGC 2011
Qy 3300 ACAGAGGGGTTCATCCACTGCATCAGATGGGGATGTACAGAGTCCAGCCACCTCTCT 3359
Db 2012 ACAGAGGGGACCTTCCACCAAGTAGAGTGGGACCTTCAGGGATCTGGGCCAACCCGTCT 2071
Qy 3360 GGTAGCACTGAGCAAGCCAGGCTGTGCTTGGGTCTGCACCCCTGAGGCGCGGTGATTC 3419
Db 2072 GAGACTTCTGGATCCGTGGCTGTCTGCAGTCTGCACACTGAAGCCCGTGCATTC 2131
Qy 3420 TCTTCCT-----GGAGCTCCAGGAACACAGGAGTGGGCTTGTCTGAGACAGTATC 3472
Db 2132 TCTCCAGGAATCAGAGCTCCAGGAACACAGGAGTGGGCTTGTCTGAGTCACT-GC 2190
Qy 3473 CTCAGGTACAGACAGGATGCACAGGTGTGCCAGGAGTGAATCTTTCCTCTGATG 3532
Db 2191 CTCAGGTACAGACAGGAGGCGCAGACAGTCCCAACACTGAAGGTTTTCCTCGGAATG 2250

Qy 3533 CACACAAAGGGCCCCACCTGCCACAGACACATAGGACTCCACAGAGTCTGGCCTCA-CC 3591
Db 2251 CACACAAAGGGCCCCACCCGCC-CAGAACAAATGGGACTCCAGAGGGCTTGGCCTCACCC 2309
Qy 3592 TCCCTACTGTCACTCTGTAGAAATCGACCTCTGCTGGCGGGTGTACCTTGA-GTACCCCT 3650
Db 2310 TCCCTATTTCTCACTCTGACCTGACCTGTGCTGGCGGGTGTACCTTGAAGTGTGCTCT 2369
Qy 3651 CTCACTTCTCTCTCAGGTTTCA-GGGACAGGCCAACCCAGAGGACAGATTTCCCTGG 3709
Db 2370 CCCATTTCTCTCTCAGGTTCTGAGGGGACAGCTGACAAGTAGGACCCGAGGACTGG 2429
Qy 3710 AGGCCACAGAGGACACCAAGAGAGATCTGTAAGTAGGCCCTTTGTAGAGTCTCCAAG 3769
Db 2430 AGGAGCATTTG-----AAGGAGAGATCTGTAAGTAAAGCTTTGTACAGGCTCCAAG 2481
Qy 3770 GTTCAG-----TTCCTAGCTGAGGCTCTCACACACTCCCTCTCTCCCCAGGCTGTGGG 3824
Db 2482 GTTCAGTTCACTTCTCACTTAAGGCTTCACACAGCTCTCTCTCTCCCCAGGCTGTGGG 2541
Qy 3825 TCTTCATTTGCCAGCTCTGCCACACTCTGCTGCTGCTGCCCTCACAGAGTCTCATGT 3884
Db 2542 TCTTCATTTGCCAGCTCTGCCGCACTCTGCTGCTGCCCTGACCCAGAGTCTCATGTC 2601
Qy 3885 CTCTTGACAGAGGAGTCTGCACTGCAAGCTTGAGGAAGCCCTTGAGGCCCAACAAGAGG 3944
Db 2602 CTCTTGACAGAGGAGTCTGCACTGCAAGCTTGAGGAAGCCCTTGAGGCCCGAGGAGG 2661
Qy 3945 CCTTGGGCTTGGTGTGTGAGGCT-----GCCACCTCTCTCT 3983
Db 2662 CCTTGGGCTTGGTGTGTCAGGCTCTCTACTAGGAGGACAGACAGGCTTCTCTCT 2721
Qy 3984 CCTCTCTCTCTCTGCGGACCTTGAGGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4043
Db 2722 CTCTACTCTAGTGGAAATTAACCTTGGGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCT 2781
Qy 4044 CCCAGAGTCTCAGGAGGCTTCCGCTTTTCCACTACATCACTTCTCAGTCTGACAGAGG 4103
Db 2782 CCCAGAGTCTCAGGAGGCTTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2841
Qy 4104 AACCCAGTGGGTTCCAGAGGCTGTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4163
Db 2842 AATCCGATGAGGCTCCAGCAACCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2901
Qy 4164 AGTCTTGTTCAGGAGTAACTCAAGAGGTCGCTGATTTGGTGGTGGTGGTGGTGGTGGT 4223
Db 2902 AGTCCGAGTTCAGAGGAGCAATCAGTAGGAGATGGTTGAGTTGGTTCATTTTCTGCTCC 2961
Qy 4224 TCAATATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4283
Db 2962 TCAAGTATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3021
Qy 4284 ATTACAGCACTGTTTCTCTGAGATCTTCGGCAAGGCTCTGAGTCTTGCAGCTGTGCT 4343
Db 3022 ATTGCCAGGACTTCTTCTCCGCTGATCTTACAGCAAGGCTCCGAGTACTTGCAGCTGTGCT 3081
Qy 4344 TTGGCATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4403
Db 3082 TTGGCATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3141
Qy 4404 TAGGCTCTCTCTATGAGGCTGCTGGGTGATTAATCAGATCATGCCCAAGAGGAGGAGGAGG 4463
Db 3142 TGGGCTCTCTCTAGGAGGCTGCTGGGCGACAAATCAGGTCATGCCCAAGAGAGGCTCC 3201
Qy 4464 TGATAATTTCTTGTCTGATGATTTGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4523
Db 3202 TGATAATTTCTTGTGCTGATTAATGCAATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3261
Qy 4524 GGGAGGAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4583
Db 3262 GGGAGGAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3321

Qy 1907 CAGAAAT-----CCCTGCTGTCAACCCACGGAAGCCACGGGAATGGCGCCAGCA 1956
Db 580 CAGAAATCCGGCTTTGCCCTCTGCAATCAACCCACGGAAGCTCCGGGAATGGCGCCAGCA 639
Qy 1957 CTCGGATCTTACGCTCCCATCCAGGCTCTGATGAGGAAGGGG-----CTTGAAC 2008
Db 640 CGCGGATCCTACGCTTCACTGTACGG-CTAAGGAGGAAGGGGTGGGTCTCGTAGT 698
Qy 2009 AGGGCTCAGGGAGCAGAGGAGGGC-----CTTACTGCGAGATGAGGAGCCCTAGA 2063
Db 699 ATGGCTTTGGGATGCAGAGAGAGGGCCAGGCTCTCTGGAAGACAGTGGAGCTCTTAG 758
Qy 2064 GGACCCAGACCTTAGA-----CACCGCACCCCTGTCTGAGACTGAGGCTGCCAC 2114
Db 759 GGACCCAGCATGCCAGGACAGGGGGCCCTGTACCCCTGTCTCAAACTGAGCCACCTTT 818
Qy 2115 TCTGGCTCAAGATCAGACGATGGGACTCAGATTCATGGGGTGGGACCCAGGCC 2174
Db 819 TCATTCAGCCAGGGAATCCTTAGGGATCAGACCCACTTCAGGGGGTTGGGCCCCAGCT 878
Qy 2175 TCGAAGGCTTACGGAGGAAGAGGAGGAGGACTCAGGGGACCTTGGAAATCCAGATCAG 2234
Db 879 GCGAGGACTCAAGGGAGGAGAGAGAGGAGGAGGACTCAGGGGACCTTGGAGTCCAGATCAG 938
Qy 2235 TGTGGACCTCGGCCCTGAGAGGTTCCAGGACAGCGTGGCCACATATGCCCCATATTTCCGT 2294
Db 939 TGGCAACCTTGG-GCTGGGGATCCTGGGCACAGTGGCCGAATGTGCCCGTGTCAATTG 997
Qy 2295 CATCTTTAGGTGAC-----AGCAGACAGCTGTGTCTCAGAAAGTGGGCGCTCAGGTCAAC 2350
Db 998 CACCTTCAAGGTTGACAGAGAGTTGAGGCTGTGTCTCAGGGCTGGGACTTCAGGTCAAC 1057
Qy 2351 AGAGGAGGAGTTCCAGGATCCATATGCCCCAAGATGTCCCCCTTCATGAGGACTGGG 2410
Db 1058 AGAGGAGGATCCAGAGATCTGCGGACCCAAAGGTGTCCCCCTTCATGAGGACTCCCC 1117
Qy 2411 ATATCCCCGGCTCAGAAAGAGGAGGACTCCACAGCTGTGGCTGTCCCCCTTTAGTAGCTC 2470
Db 1118 ATACCCCGGGCCAGAAAGAGGATGCCACAGATCTGGAATTAATTTCTTAGCTC 1177
Qy 2471 TAGGGGACACAGATCAGGATGGGGTATGTTCATTTCTCAGCTGTACACAGGCGAA 2530
Db 1178 TGGGGAAACCTGATCAGGGATGGCCCTAAGTGACAATCTCATTTGTACACAGCAGAG 1237
Qy 2531 GTTGGGGGCCCTCAGGAGATGGGTCTTGGGGTAAAGGGGAGTGTCTACTCATGTCA 2590
Db 1238 GTTGGGAACCTCAGGAGATAGGTGTGTGTAAAGAGAGCTGTCTGCTCATTTCA 1297
Qy 2591 GGGAAATGGGGTTGAGGAAGCAGAGGCGCTGGCAGGAATAAAGATGAGTGAGACAGACA 2650
Db 1298 GGGGTTCCCCCTTGAGAAAGGCGAGTCCCTGGCAGGAGTAAAGATGATTAACCCACAGG 1357
Qy 2651 AGGCTATTTGGAAATCCACACCCAGCAACAAAGGGGTAGCCCTGGACACCTCACCCAGGA 2710
Db 1358 AGGCAATCATACGCTTACCCCTAGAACCAAGGGGTAGCCCTGGACACCGCATGGGG 1417
Qy 2711 T-----GTGGCTTTTCTTTTCACTCTCTTCCAGATCTGGGCGAGGTGAGGACCTCAT 2763
Db 1418 TAACAGGATGTGGCCCTCTCTACTTTCTTCTCCAGATCTCAGGGAGTTGATGACCTTGT 1477
Qy 2764 TCTCAGAGGGTGACTCAGGTCAAGCTAGGAGCCGCCATCTGTGCTTAAAGACAGAGCGCTC 2823
Db 1478 TTTCAGAAAGGTGACTCAGTCAACACAGGGG---CCCCCTCTGTGCGACAGATGCACTGGTT 1534
Qy 2824 CCAGATCTGCCATCGCTTTCGGGTGAGGAACATAGGAGGAGCTAGGGTACCCAGGAC 2883
Db 1535 CTAGGATCTGCAAGCATCTCAGGTGTGAGAGGCTGAGGTAGGATGAGGAGTACCCCTGGGC 1594
Qy 2884 CAGAACTGT-AGGAGAGACTGCACAGAAATCAGCCCTGCCCTGTCTACCCACAGAGAG 2942
Db 1595 CAGAAATGCAAGAGGGGGCCCATAGAAATCTGCCCTGCCCTGCGGTACTTTCAGAGAC 1654

Qy 2943 CATGGGCTGGGCGCTCTCCGAGGTCCTTCCGTTATCTCTGGATCATTTGATGTCAAGGAC 3002
Db 1655 CTGGGCAAGGCTGTGACGTGAAGTCCCTCCATTAT-CTGGGATCTTTGATGTCAAGGAA 1713
Qy 3003 GGGAGGCTTTGGTCTGAGAGGCTGCCCTCAGTCACTAGAGGAGCGTCCCGAGCCCT 3062
Db 1714 GGGAGGCTTTGGTCTGAAGGGCTTGGAGTCAAGTCACTAGAGGAGGCTCTCAGCCCT 1773
Qy 3063 GCCAGGATCAAGGTGAGGACCAAGCGGACCTCACCCAGGACACATTAATTTCAATGA 3122
Db 1774 GCCAGGATGACGTGAGGACCAAGCGGACTCGTCAACCCAGGACACTTGGATCTCAATGA 1833
Qy 3123 ATTTGATATCTTTGCTGCCCTTCCCTCCCAAGACCTAGGACGTGTGGCCAGATGTTGT 3182
Db 1834 A-TTTGACATCTCTGTGTCTTCCGAGGACCTGTGTACGTATGGCCAGATGTGGT 1892
Qy 3183 CCCCTC-CTGTCTTCCATTCTTATCATGATGATGAACTCTTG--ATTGATTTCTCA 3239
Db 1893 CCCCTCTATCTCTGTACCATATCAGGGATGTGAGTCTTGACATGAGAGATTTCTCA 1952
Qy 3240 GACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCCGTAGA 3299
Db 1953 AGCCAGCAAAAGGGTGGAT-TAGGCCCTACAGGAGAAAGGTGAGGGCCCTGAGTGAGC 2011
Qy 3300 ACAGAGGGGTATCTCCACTGATGAGTGGGATGTCAAGAGTCCAGCCACCCCTCT 3359
Db 2012 ACAGAGGGGACCTCCACCCCAAGTAGAGTGGGACCTCACGGAGTCTGGCCAACTGCT 2071
Qy 3360 GGTAGCACTGAGAACCCAGGCTGTCTTGGGCTGCACCTGAGGGCCCTGATTTCC 3419
Db 2072 GAGACTTCTGGGAATCCCTGTCTGTGCTGTGACACTGAAGGCCCTGCGATTTCC 2131
Qy 3420 TCTTCTCT-----GGAGCTCCAGGAACAGGCACTGAGGCTTGGTGTGACACAGTATC 3472
Db 2132 TCTCCAGGATCAGGACTCCAGGAACAGGCACTGAGGCTTGGTGTGAGTCACT-GC 2190
Qy 3473 CTCAGGTCAAGACAGAGGATGACAGGGTGTGCCAGCAGTGAATGTTGCCCTGAATG 3532
Db 2191 CTCAGGTCAAGACAGAGGAGGAGCAGACAGATGCCAACACTGAAGGTTTGGCTGGAATG 2250
Qy 3533 CACACCAAGGGCCCACTCCACAGGACACATAGGACTCCACAGACTGTGGCCCTCA-CC 3591
Db 2251 CACACCAAGGGCCCACTCCAGGAC-AGAACAAATGGGACTCCAGAGGGCTGGCCCTCACCC 2309
Qy 3592 TCCCTACTGTCTAGTCTAGAACTGACCTGTCTGGCGGCTGTACCTTGA-GTACCTCT 3650
Db 2310 TCCCTATTTCTAGTCTGACCTGAGCATGTCTGGCGGCTGTACCTGAGGTGCCCT 2369
Qy 3651 CTCACTTCTCTTCTAGGTTTCA-GGGGACAGGCCAACCCAGAGGACAGGATTCCTCTGG 3709
Db 2370 CCCACTTCTCTTCTAGGTTCTGAGGGGACAGGCTGACAAGTAGGACCCGAGGCACTGG 2429
Qy 3710 AGGCCACAGAGGACCAAGGAGAGATCTGTAAGTAGGCTTTGTTAGAGTCTTCCAG 3769
Db 2430 AGGACATTG-----AAGGAGAAAGATCTGTAAGTAAGCTTTGTGACAGCTTCCAG 2481
Qy 3770 GTTCAG-----TCTCAGCTGAGGCTCTCACACTCCCTCTCTCCCGAGGCTGTGGG 3824
Db 2482 GTTCAGTTCTCAGTTCTCAGCTAAGGCTCACACGCTCTCTCTCCCGAGGCTGTGGG 2541
Qy 3825 TCTTCATTGCCAGGCTCTCCACACTCTCTGCTGTCTGCTGTGACGAGAGTCACTATGT 3884
Db 2542 TCTTCATTGCCAGGCTCTCTCCCGACTCTCTGCTGTCTGCTGTGACGAGAGTCACTATG 2601
Qy 3885 CTCCTGACAGAGGCTGTGACTGCAAGCTGAGGAGGCTTGTAGGCGCCCAAGAGG 3944
Db 2602 CTCCTGACAGAGGCTGAGGCTGCAAGCTGCAAGGCTGAAAGGCTTGTAGGCGGAGGAGG 2661
Qy 3945 CCCTGGGCTGTGTGTGCGAGGCT-----GCCACTCTCTCT 3983
Db 2662 CCCTGGGCTGTGTGTGCGGAGGCTCTCTGCTACTGAGGAGCAGGACGCTTCTTCTCT 2721
Qy 3984 CCTCTCTCTCTGCTCTGGGCACTCTGGAGGAGGTGCCCACTGCTGGGTCAACAGATCCTC 4043

Db	2722	CTTCTACTAGTGGAAAGTTACCCCTGGGGAGGTGCCTGCTGGCGACTCACCGAGTCCCTC	2781
Qy	4044	CCCAGAGTCTCTCAGGAGGCTCCGGCTTTTCCACTTACCATCAACTTCCACTCGACAGAGGC	4103
Db	2782	CCCACAGTCTCTCAGGAGGCTCCAGCTTCTCGACTACCATCACTACACTCTTTGGAGAC	2841
Qy	4104	AACCCAGTGAAGGTTCCACAGACCGGTGAAGAGAGGGGCCAAGCACCTCTGTATCCTGG	4163
Db	2842	AATCCGATGAGGCTCCAGCAACCAAGAAGAGGAGGGGCCAAGAAATGTTTCCCGACCTGG	2901
Qy	4164	AGTCCCTGTTCGAGCAGTAATCACTAAGAAGTGGCTGATTTGGTTGGTTTCTGCTCC	4223
Db	2902	AGTCCGAGTTCCAAGCAGCAATCACTAGGAAGATGGTTGAGTTGGTTCATTTTCTGCTCC	2961
Qy	4224	TCAAATATCGAGCCAGGAGCCAGTCACAAAAGGAGAAATGCTGGAGAGTGTCTATCAAAA	4283
Db	2962	TCAAGTATCGAGCCAGGAGCCGGTCACAAAAGGAGAAATGCTGGAGAGTGTCTCAGAA	3021
Qy	4284	ATTACAAGCACTGTTTTTCTAGATCTTCGGGAAAGCCTCTGAGTCTTCGAGCTGGTCT	4343
Db	3022	ATTGCCAGGACTCTTTTCCCGTGATCTTCAGCAAAAGCTCCGAGTACTTCAGCTGGTCT	3081
Qy	4344	TTGGCATTGACGTGAAGGAGCAGACCCACCGGCCACTCCTATGTCCTTGTCACTGCC	4403
Db	3082	TTGGCATCGAGGTGGTGAAGTGGTCCCACTCAGCCACTGTGATCCTTGTCACTGGC	3141
Qy	4404	TAGGTCCTCTATGATGSCCTGCTGGGTGATATACAGATCATGCCCAAGACAGGCTTC	4463
Db	3142	TGGCCCTCTCTACGATGGCTGCTTGGGGACATCAGGTCATGCCCAAGACAGGCTTC	3201
Qy	4464	TGATAATTGTCTGTGTCATGATTGCAATGGAGGGCGGCCATGCTCTCGAGGAGGAATCT	4523
Db	3202	TGATAATCGTCTGGCCATAATCGCAATAGAGGGCGACTGTGCCCCCTGAGGAGAAATCT	3261
Qy	4524	GGAGAGCTGAGTGTGATGGAGGTGTATGATGGGAGGGACACAGTGCCTATGGGGAGC	4583
Db	3262	GGAGAGCTGAGTATGTTGGAGGTGTTTGAGGGGAGGGAGACAGTGTCTTCGCACATC	3321
Qy	4584	CCAGGAAGCTGCTCACCCAAAGATTTGGTGCAGGAAAAGTACCTGGAGTA-CGGCAGGTGC	4642
Db	3322	CCAGGAAGCTGCTCATGCAAGATCTGGTGCAGGAAAACACTACCTGGAGTACCGCAGGTGC	3381
Qy	4643	CGACAGTGATCCCGCACGCTATAGTTCCTGTGGGTTCGAAGGGCCCTCGCTGAAACCA	4702
Db	3382	CCGGCAGTGATCCTGCATGCTACGAGTTCCTGTGGGTTCGAAGGGCCCTCATTTGAAACCA	3441
Qy	4703	GCTATGTGAAGTCCCTTGAGTATGTATCAAGTTCAGTGCAGAGAGTTCGCTTTTCTTCC	4762
Db	3442	GCTATGTGAAGTCCCTGACCATACACTAAAGATCGGTGGAGAACCTTCACATTTCTTACC	3501
Qy	4763	CATCCCTCGGTGAAGCAGCTTTTGAGAGAGGAGGAAGGGAGTCTGAGCATGAGTTGTCAG	4822
Db	3502	CACCCCTGCATGACGGGCTTTTGAGAGAGGAGNAGAGTGAAGTCTCAGCATGTTGTCAG	3561
Qy	4823	CCAAGGCCAGTGGGAGGGGACTGGGCCAGTGCACCTTCCAGGCCCCGCTCCACGAGTT	4882
Db	3562	CCAGGGCCAGTGGGAGGGGTCTGGGCCAGTGCACCTTCCAGGGCCCCCATTTAGCTT	3621
Qy	4883	CCCTCGCTCGTGTGACATGAGGCCCATTCIT--CACTCTGAAGAGAGCGGTCAAGTGTTC	4940
Db	3622	CCACTGCCCTCGTGATATGAGGCCCATTCCTGCCTCTTTGAAGAGAGCAGTCAGCATTC	3681
Qy	4941	TCAGTAGTAGGTTTCTGTCTATTGGGTGACTTTGGAGATTTATCTTTGTCTCTTTTGGG	5000
Db	3682	TTAGCAGTGAGTTCTGTCTCTGGATGACTTTTGAGATTTATCTTTCTTCTCTGTGGA	3741
Qy	5001	ATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACCTTCAGCATCCAAGTTTATGA	5060
Db	3742	ATTGTTCAAATG-TTCCCTTTTAAACAATGGTTGGATGAACCTTCAGCATCCAAGTTTATGA	3800
Qy	5061	ATGACAGCAGTCACAC--AGTTCTCTGTATATAGTTTAAAGGTAAGAGTCTTGTTTATA	5118

D	b	3801	-ATGACAGTGTAGTCACACATAGTGCTCTTTATATATAGTTTTAGGGGTAAAGAGTCCTCTGTTTTTA	3866
Q	y	5119	TTCAGATTGGGAATCCCATTTCTATTTCTGAAATTG--GGATAAATAACAGCAGTGGGAATA	5176
D	b	3861	TTCAGATTGGGAATCCCATTTCTGAAATTG--GGATAAATAACAGCAGTGGGAATA	3920
Q	y	5177	GTAATTGGGAATCCCATTTCTGAAATTG--GGATAAATAACAGCAGTGGGAATA	5232
D	b	3921	GTAATTGGGAATCCCATTTCTGAAATTG--GGATAAATAACAGCAGTGGGAATA	3974
Q	y	5233	ATTAAGAGTAGTCAATCTTGCCTTTATACCTCAGTCTATTTCTTAAAATTTTAAAGAT	5292
D	b	3974	-TCAAAGATAGTAAATTTTGCCTTTATACCTCAGTCTATTTCTTAAAATTTTAAAGAT	4029
Q	y	5293	ATATGATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAAAATTAATCTGAATA	5352
D	b	4030	ATGCTGATGTT-----TTTGCTTCTTTGAGAATGCAAAAGAAAAATTAATCTGAATA	4080
Q	y	5353	AAGAATCTTCTGCTGACTGGCTCTTTTCTTCCATGCACTGAGCATCTGCTTTTGG	5412
D	b	4081	---AATCTTCTGCTGACTGGCTCTTTTCTTCCATGCACTGAGCATCTGCTTTTGG	4137
Q	y	5413	AAGGCCCTCGS	5422
D	b	4138	AAGGCCCTCGS	4147
R	E	S	RESULT 11	
X	I	D	X26974 standard: cDNA; 4204 BP.	
A	C	AC	X26974;	
D	T	DT	25-JUN-1999 (first entry)	
D	E	DE	cDNA encoding MAGE-3 polypeptide.	
K	W	KW	MAGE-3 tumour associated gene; human leucocyte antigen Class II;	
K	W	KW	autologous CD4+ cell; MAGE-3 related disease; cancer; melanoma;	
K	W	KW	osteosarcoma; leukemia; carcinoma; ss.	
O	S	OS	Homo sapiens.	
F	H	FH	Key	
F	T	FT	Location/Qualifiers	
F	T	FT	: 2465..3409	
F	T	FT	/tag= a	
F	T	FT	/product= "MAGE-3"	
P	N	PN	WO9914326-A1.	
P	D	PD	25-MAR-1999.	
P	F	PF	04-SEP-1998; U18601.	
P	R	PR	12-SEP-1997; US-928615.	
P	A	PA	(LUDW-) LUDWIG INST CANCER RES.	
P	A	PA	(UYVR-) UNIV VRIJE BRUSSEL	
P	I	PI	Boon-Faller T; Chaux P, Corthals J, Heirman C,	
P	I	PI	Luiten R, Stroobant V, Thielemans K, Van Der Bruggen P;	
P	R	PR	WPI; 99-244031/20.	
D	R	DR	p-PDSR; Y01720.	
P	T	PT	Isolated peptides that bind to human leucocyte antigen class II	
P	T	PT	molecules	
P	S	PS	Example 5; Page 65-67; 88pp; English.	
C	C	CC	The present sequence represents the MAGE-3 tumour associated gene.	
C	C	CC	Peptides that bind human leucocyte antigen (HLA) Class II molecules	
C	C	CC	can be derived from the MAGE-3 protein. These peptides and	
C	C	CC	autologous CD4+ cells that bind to a complex of MAGE-3 peptide	
C	C	CC	and HLA Class II, are used to treat MAGE-3 related diseases,	
C	C	CC	particularly cancers (e.g. melanoma, osteosarcoma, leukemia and	
C	C	CC	various forms of carcinoma). The peptides are also used to produce	
C	C	CC	specific antibodies. Detection of the peptides, e.g. in binding	
C	C	CC	assays, particularly with antibodies, is used for diagnosis of such	
C	C	CC	diseases.	
S	Q	SQ	Sequence 4204 BP; 944 A; 1144 C; 1223 G; 893 T;	
Q	U	QU	Query Match 31.1%; Score 1762.6; DB 1; Length 4204;	
B	E	BE	Best Local Similarity 81.0%; Pred. No. 0;	
M	A	MA	Matches 2315; Conservative 0; Mismatches 474; Indels 68; Gaps	
Q	Y	QY	2790 AGGACCCCCCATCTGGTCTTAACAGACAGCGGTCCCAGGATCTCCCATGCTTCGGGTGA	2849

Db	1365	AGGGGCCCTTATGTGTGGACAGATGCACTGTGCTTAGATCTGCCAAAGCATCCAGGTGA	1424
Qy	2850	GGAACTATGAGGGAGGACTGAGGTTACCCAGACACCAAGCACTGA-GGGAGACTGCAAG	2908
Db	1425	AGAGACTGAGGGAGGATTGAGGTACCCCTGGACAGAAATGCGGACTGGGGCCCATAA	1484
Qy	2909	AAATAGCCCTGCCCTGCTGTCTACCCAGAGAGCATGGGCTTGGGCGGTCTGCGGAGGT	2968
Db	1485	AAATCGCCCTGCTCTGCTGTATCTCAGAGAGCTGGCAGGGCTGTACGCTGAGGTC	1544
Qy	2969	CTTCCGTTATCCTTGGGATCATGTAGTCAAGGACAGGGGAGGCCCTTGGTCTGAGAGGCTG	3028
Db	1545	CCTCCATATCTTAGAGTCACTGATGTCAAGGAAGGGGAAGCCTTGGTGTAGAGGGGCTG	1604
Qy	3029	CGCTCAGGTCAGTAGAGGAGCGTCCAGGCCCTGCCAGGAGTCAAGGTGAGGACCAAGC	3088
Db	1605	CACTCAGGCAAGTAGAGGAGGCTCTCAGACCCTACTAGGAGTGGAGGTGAGGACCAAGC	1664
Qy	3089	GGGCACCTCACCCAGGACACATAATTCCAATGAATTTTGATATCTCTTGTG-CCTTTC	3147
Db	1665	AGTCTCCTCACCCAGGATACATGGACTTCAATAAATTTGGACATCTCTCGTTGTCTTTC	1724
Qy	3148	CCNAGGACCTAGGCAGCTGTGGCCAGATGTTTGTCCCTCCTGTCTTCCATTCCTTAT	3207
Db	1725	CGGAGGACCTGGGAATGATGGCCAGATGTGGGTCCCTCATGTGTTTCTGTACCATAT	1784
Qy	3208	CATGGATGTCAACTCTTG--ATTTGGATTCTCACACAGCAAAAGGCGAGGATCCAGGC	3265
Db	1785	CAGGTATGTGAGTCTTGACATGAGAGATTCTCAGGCCACAGAAAGGAGGGGAT-TAGGC	1843
Qy	3266	CCTGCCAGAAAATAAAGGCCCTTGCCTGAGAACAGAGGGGCTCATCTCAGTCACTGAG	3325
Db	1844	CCTAATAGGAAAGGTGAGGGCCCTGAGTGAGCACAGAGGGGATCCTCCACCCCACTAG	1903
Qy	3326	AGTGGGATGTCACAGAGTCCAGCCCAACCTCCTCGTAGCTAGAAAGCAGAGGGCTGTG	3385
Db	1904	AGTGGGACCTCACAGAGTCTGSCCAACCTCCTCAGCAGTTCTGGGAATCCGTGGCTGCG	1963
Qy	3386	CTTGGGCTGTCACTCAGGAGCCGCTGGATCTTCCTC-----CTGGAGCTCCAGGA	3438
Db	1964	TTTGTGCTGTGCACATTTGGGGCCCGTGATTTCTCTCCAGGAATCAGGAGCTCCAGGA	2023
Qy	3439	ACCAGGCAGTGAGGCCCTTGGCTGTGACAGAGTATCCTCAGGTCACAGAGCAGAGGATGCAC	3498
Db	2024	ACAAGGCAGTGAGGACTTGGTCTGAGGCAGTGTCTCCTCAGGTCACAGAGTAGAGGGGCTC	2083
Qy	3499	AGGTTGCCAGCAGTGAATGTTTGGCTTGAATGCACACCAAGGGCCCCACCTGCCACAG	3558
Db	2084	AGATAGTGCCACGGTGAAGGTTTGGCTTTGGATTCAAAACCAAGGGCCCCACCTGCCCCAG	2143
Qy	3559	GACACATAGGACTCCACAGAGTCTGGCTCA-CCTCCCTACTGTTCAGTCCGTAGATCG	3617
Db	2144	AACACAT-GGACTTCCAGAGGCGCTGGCTCACCCTCAATACTTTCAGTCTGTGACGCTCA	2202
Qy	3618	ACCTCTGCTGGCCGCGTACCCCTGA-GTACCCCTCTCACTTCCCTCCTTCAGGTTTTCAGG	3676
Db	2203	GCATGCGCTGGCCGGATGTACCCCTGAGGTGCGCTCTCACTTCCCTCTCAGTTCTGAGG	2262
Qy	3677	GGACAGGCCAACCCAGGACAGGATTCCTCTGGAGGCCACAGAGAGAGCAACAGGAGAAG	3736
Db	2263	GGACAGGCTGACCTGGAGGACACAGAGGCCCGCCGGAGGACACTG-----AAGGAGAAG	2315
Qy	3737	ATCTGTAAGTAGGCCCTTGTAGAGTCTCCAGG-----TTACGTTCTCAGCTGAGGCCCT	3791
Db	2316	ATCTGTAAGTAAGCCCTTGTTAGCGCTTCCAAGTTTCCATTCACTGACTACTCAGCTGAGGTCT	2375
Qy	3792	CTCACACACTCCCTCTCTCCCAAGGCCGTGTGGGTCTTCATTGGCCCAAGCTCCTGTGCCACAC	3851
Db	2376	CTCACATGCTCCCTCTCTCCCAAGGCCAGTGGGTCTCCATTGCCCAAGCTCTGTGCCACAC	2435
Qy	3852	TCCTGCGCTGCCCTGACGAGAGTCATCATGTCTCTTTGACGAGAGGAGTCTGCACGTCA	3911
Db	2436	TCCGCGCTGTTGGCCCTGACACAGTCACTATGCCCTTTTGGACAGAGAGTCTGACACTGCA	2495

Qy 4968 TGACCTGGAGATTTATCTTGTCTCTTTTGGAAATGTTCAAAATGTTTTTTTTTAAAGGA 5027
Dy 3576 TGACCTTTCAGATTAATCTTGTCTCTCTGTTGAGTGTGTTCAAAATG-TTCCCTTTTAAACGA 3634
Qy 5028 TGGTTGAATGAACCTCAGCATCCAGTTTATGAATGACAGAGTGCACAC--AGTTCTGTG 5085
Dy 3635 TGGTTGAATGAGCGTCAGCATCCAGTTTATGAATGACAGTGTGCACATATGCTGTG 3694
Qy 5086 TATATAGTTTAAAGGTAAGAGCTCTGTTTATTCAGATGGGAAATCCATCTTATTTT 5145
Dy 3695 TATATAGTTTAAAGGTAAGAGCTCTGTTTATTCAGATGGGAAATCCATCTTATTTT 3754
Qy 5146 GTGAATTTGG--ATAATACAGCAGTGGAAATAAGTA----CTTAGAATGTGA-AAAAT 5198
Dy 3755 GTGAATTTGCACATATAATAGCAGTGGTAAAGATTTTCTTAAATTTGTGAGCGAAT 3814
Qy 5199 AGCAGTAAATAGATGAGATAAGAACTAAAGAAATTAAGAGATAGTCAATTTTCCTT 5258
Dy 3815 AGCAATAACATACATGAGAT----AACTCAAGAAATCAAAAGATAGTGTATTTGCTT 3870
Qy 5259 ATACCTCAGTCTATCTGTAAATTTTAAAGATATATGCATACCTGGATTTTCCTTGGCT 5318
Dy 3871 GTACCTCAATCTATCTGTAAATTTTAAAGATATATGCATACCTGGATTTTCCTTGGCT 3927
Qy 5319 TCTTTGAGATGTAGAGAAATTAATCTGAATAAGAAATTTCTCTGTACTGGCTT 5378
Dy 3928 TCTTTGAGATGTAGAGAAATTAATCTGAATAAGAAATTTCTCTGTACTGGCTT 3987
Qy 5379 TTTCTTCTCCAGTCTGAGCATCTGCTTTTGGAGGCTTGGTGTAGTGTAGTGTAGGATG 5438
Dy 3988 TTTCTTCTCCAGTCTGAGCATCTGCTTTTGGAGGCTTGGTGTAGTGTAGTGTAGGATG 4047
Qy 5439 CTAAGGTAAAGCAGACTCATACCCACCATAGGCTGCTAGAGTCTAGGAGTGCAGTGCAC 5498
Dy 4048 CTAAGGTAAAGCAGACTCATACCCACCATAGGCTGCTAGAGTCTAGGAGTGCAGTGCAC 4107
Qy 5499 GTAATCGAGTGGCAGATGCTCTCTAAAGATGTAGGAGAAATGAGAGAGGGGTGAGGG 5558
Dy 4108 GTAATCGAGTGGCAGATGCTCTCTAAAGATGTAGGAGAAATGAGAGAGGGGTGAGGG 4167
Qy 5559 TGTGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATGCC 5595
Dy 4168 TGTGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATGCC 4204

RESULT 12

ID Q72482 standard; DNA: 2531 BP.
AC Q72482;
DT 22-JUN-1995 (first entry)
DE Tumour rejection antigen MAGE-4 gene.
KW Tumour rejection antigen; melanoma antigen-4; MAGE-4; MAGE-3;
KW cancer; cytolytic T cells; antigen D; human leucocyte antigen;
OS ss.
FH Homo sapiens.
FT key Location/Qualifiers
FT cds 625..1578
FT /*tag= a
PN W09423031-A.
PD 13-OCT-1994.
PF 17-MAR-1994; U02877.
PR 26-MAR-1993; US-037230.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon-fallier T. Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;
DR WPI: 94-333192/41.
PT New tumour rejection antigen precursor MAGE3 - useful in
PT treatment and diagnosis of cancer
PS Example 30; Page 66; 105pp; English.
CC Q72482 is the gene which contains the coding sequence for melanoma
CC antigen-4 (MAGE-4). Another melanoma antigen MAGE-3 is encoded by
CC Q72470, this is a tumour rejection antigen precursor. Melanoma
CC characterised by the expression of MAGE-3 can be detected, or
CC monitored, by contacting a test sample with an agent that can

CC recognise MAGE-3. The melanoma can be treated by the administration
CC of cytolytic T cells specific for the complex of antigen D (the
CC mature rejection antigen derived from MAGE-3) and a human leucocyte
CC antigen (esp. HLA-A1). 608 A; 593 C; 700 G; 630 T;
SQ Sequence 2531 BP;

Query Match 29.6%; Score 1678.8; DB 1; Length 2531;

Best Local Similarity 84.4%; Pred. No. 0;

Matches 2074; Conservative 0; Mismatches 317; Indels 65; Gaps 14;

Qy 3256 GGATCCAGCCCTGCCAGGAAAAATATAAGGCCCTGGTGAGACACAGAGGGGTCTATCC 3315

Dy 1 GGATCCAGCCCTGGCTGGAGAAATGTGAGGCCCTGAGTGAACACAGTGGGATCATCC 60

Qy 3316 ACTGCATCAGAGTGGGATGTACAGAGTCCAGGCCCTCTCTGTAGCACTGAGAAGC 3375

Dy 61 ACTGCATCAGAGTGGGATGTACAGAGTCCAGGCCCTCTCTGTAGCACTGAGAAGC 120

Qy 3376 CAGGCTGTGCTTGGGCTGTGCACCCCTGAGGCCCTGGATTCCTTCTCTGGAGCTCCA 3435

Dy 121 CAGGCTGTGCTTGGGCTGTGCACCCCTGAGGCCCTGGATTCCTTCTCTGGAGCTCCA 180

Qy 3436 GGAACAGGCTGAGGCTTGGTCTGAGACAGTATCCTCAGGTACAGAGCAGAGGATG 3495

Dy 181 GGAACAGGCTGAGGCTTGGTCTGAGACAGTATCCTCAGGTACAGAGCAGAGGATG 240

Qy 3496 CACAGGGTGTCCAGCAGTGAATGTTCCCTGATGACACCAAGGCCCTCCACCTGCCA 3555

Dy 241 CACAGGGTGTCCAGCAGTGAATGTTCCCTGATGACACCAAGGCCCTCCACCTGCCA 300

Qy 3556 CAGGACACATAGGACTCCACAGAGTCTGGCCTCACTTCTCTCTCTCTCTCTCTCT 3615

Dy 301 CAGGACACATAGGACTCCACAGAGTCTGGCCTCACTTCTCTCTCTCTCTCTCTCT 360

Qy 3616 CGACCTCTGCTGGCGGTGTACCTCTGA-GRACCTCTCACTTCTCTCTCTCTCTCTCT 3674

Dy 361 CGACCTCTGCTGGCGGTGTACCTCTGA-GRACCTCTCACTTCTCTCTCTCTCTCTCT 420

Qy 3675 GGGGACAGCCCAACCCAGAGGACAGGATTCCTGAGGCCACAGAGGACCAAGGAGA 3734

Dy 421 GGGGACAGCCCAACCCAGAGGACAGGATTCCTGAGGCCACAGAGGACCAAGGAGA 478

Qy 3735 AGATCTGTAAGTAGGCTTTGTTAGAGTCTCAAGGTTCTCAGTCTCAGTCTCAGTCTCT 3794

Dy 479 AGATCTGTAAGTAGGCTTTGTTAGAGTCTCAAGGTTCTCAGTCTCAGTCTCAGTCTCT 538

Qy 3795 ACACACTCCCTCTCTCCAGGCTGTGGGTCTTCATTGCCAGCTCTCTGCCACCTCC 3854

Dy 539 ACATGCTCCCTCTCTCCAGGCTGTGGGTCTTCATTGCCAGCTCTCTGCCACCTCC 598

Qy 3855 TGCCTGCTCCCTCTGACGAGATCATGCTCTCTGAGCAGAGGATCTGCACTGCAAGC 3914

Dy 599 TGCCTGCTCCCTCTGACGAGATCATGCTCTCTGAGCAGAGGATCTGCACTGCAAGC 658

Qy 3915 CTGAGGAAAGCCCTCTGAGGCCCAACAGAGGCCCTGGGCTGTGTGTGTCAGGCTGCCA 3974

Dy 659 CTGAGGAAAGCCCTCTGAGGCCCAACAGAGGCCCTGGGCTGTGTGTGTCAGGCTGCCA 718

Qy 3975 C-----CTCCT 4010

Dy 719 CT 778

Qy 4011 AGGAGGTCCT 4070

Dy 779 AGGAGGTCCT 838

Qy 4071 TTTCCACTTACCATCACTTCACTTCACAGAGGCAACCCAGTGTGGTTCACAGCCGCTG 4130

Dy 839 TTTCCACTTACCATCACTTCACTTCACAGAGGCAACCCAGTGTGGTTCACAGCCGCTG 898

Qy 4131 AAGAGGAGGGGCAAGCAGCT 4190

Dy 4131 AAGAGGAGGGGCAAGCAGCT 4190

RESULT 13

Q72405
ID: 072483 standard: DNA: 2531 BP

DT 22-JUN-1995 (first entry)

KW Tumour rejection antigen; melanoma antigen-41; MAGE-41; MAGE-3;

KW SS.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT - /*tag= a

PD 13-OCT-1994 .

PR 26-MAR-1993; US-037230.

PI BOON-falleur 1, Gaugler B, van DEN EYNDE B, van DER BRUGGEN
DR WPT. 04'-323102/11

PT
FI
New tumour rejection antigen precursor MAGE3 - useful in treatment and diagnosis of cancer

DISCLOSURE, Page 00, 103pp, English.

CC 072470, this is a tumour rejection antigen precursor Melanomas

CC monitored, by contacting a test sample with an agent that can

CC of cytolytic T cells specific for the complex of antigen D (the

CC antigen (esp. HLA-A1).

Best Local Similarity 84.4%; Pred. No. 0;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db	1139	CCACGACACACTACACCCCTTGTCACCTGGCGCTTCTCTATGATGGCCTGCTGG	1198
Qy	4431	GTGATATCAGATCATGCCCAAGACAGGCTTCCCTCATATTTGTCTCGTCATGATTCGAA	4490
Db	1199	GTATATCATGATCTTCCCAAGACAGGCTTCTCATATCGTCTGGCACAATTCGAA	1258
Qy	4491	TGAGGGCGCCCATGCTCTCTGAGGAGAAATCTGGAGAGCTGAGTGTGATGGAGTGT	4550
Db	1259	TGAGGGCGACAGCGCTCTGAGGAGGAAATCTGGAGGAGCTGGGTGTGATGGGGTGT	1318
Qy	4551	ATGATGGGAGGAGCACATGCTATGCGGAGCCCGAGGAAGCTGCTCACCCAGAATTTGG	4610
Db	1319	ATGATGGGAGGAGCACATGCTATGCGGAGCCCGAGGAAGCTGCTCACCCAGAATTTGG	1378
Qy	4611	TGCAGGAAGAAGTACCTGGAGTA-CGGCAGGTCCGCGACAGTATCCCGACGCTATGAGT	4669
Db	1379	TGCGAGGAATACTACCTTGGAGTACCGCGAGGTACCCGCGAGTAATCTCGCGCGCTATGAGT	1438
Qy	4670	TCCTGTGGGTCCTCAAGGGCCCTCGCTGAAACAGACTATGTGAAAGTCCCTTGAGTATGTA	4729
Db	1439	TCCTGTGGGTCCTCAAGGGCTCGCTGCTGAAACAGACTATGTGAAAGTCCCTTGAGCATGTGG	1498
Qy	4730	TCAGGTCTAGTGCAGAGTTTCGTTTTCTTCCCATCCCTCGGTGAAGCAGCTTTGAGAG	4789
Db	1499	TCAGGTCTAATGCAAGAGTTTCGATTCGCTACCCATCCCTCGGTGAAGCAGCTTTGTTAG	1558
Qy	4790	AGGAGGAAGGAGGAGTCTGAGCATGAGTGCAGCCAGCCAGCGGAGTGGGAGG- ---GGACT	4845
Db	1559	AGGAGGAAGGAGGAGTCTGAGCATGAGTGCAGCCAGCGGCTGTGGGGAAGGGGCGGCT	1618
Qy	4846	GGCCAGTGCACCTCCAGGGCGGGTCCAGCAGCTTCCCTGCTCGTGTGACATGAGG	4905
Db	1619	GGCCAGTGCATCTACA- -GCCCTGCGAGCAGCTTCCCTTGCTGTAAACATGAGG	1676
Qy	4906	CCCATCTTTCACATC- ---TGAAAGAGCGGTCAGTGTCTTCAGTAGTAGTTCCTGTTCT	4961
Db	1677	CCCATCTTTCACATCTCTTTGAAGAAATAGTCAGTGTCTCTTAGTAGTGGGTCTCTATTTT	1736
Qy	4962	ATTGGGTGACTTCGAGATTTACTTTGTCTCTTTTGGAAATGTTCAAATCTTTTCTTTT	5021
Db	1737	GTGTGATGACTTCGGAGATTATCTCTGTTTCTCTTTTGAATGTTCAAATG-TTCCCTTTT	1795
Qy	5022	AAGGGATGTTGAATGAACCTTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTC	5081
Db	1796	ATGGATGGTGAATTAACCTCAGCATCCAAGTTTATGAATCGTAGTTAAAGTATATTCG	1855
Qy	5082	TGTGTATATAGTTTAAGGTAAGAGTCTTGTGTTTTTATTCAGATTGGGAAATCCCATCTTA	5141
Db	1856	TGTTAATATAGTTTAGGATGAAGTCTTGTTTTTTATTTCAGATTGGGAAATCCGTTCTTA	1915
Qy	5142	TTTTGTGAATTTGGG- ---ATAATAACAGCAGTGGGAATAGTACTTAGAATGTGAAAAATG	5198
Db	1916	TTTTGTGAATTTTGGGACATAATAACAGCAGTGGGATAGTATTTAGAAGTGTG- ---AAT	1972
Qy	5199	AGCAGTAAAAATAGATCAGATATAAGAACTAAAGAAATTAAGAGATAGTCAATTTCTGCCTT	5258
Db	1973	CACCGTGAATAGCTGAGAT- ---AATTAAGATACCTTAATTCGCGCTT	2020
Qy	5259	ATACCTCAGTCTATTCTGTAATAATTTTAAAGATATATGCATACCTGGATTTCTTTGGCT	5318
Db	2021	ATGCCTCAGTCTATTCTGTAAAAATTTAAAAATATATATGCATACCTGGATTTCTTTGGCT	2080
Qy	5319	TCATTGAGATGTAAGAGAAATTAATCTGAATAAAGAAATTTCTTCTCTGTCACGTGGCTCT	5378
Db	2081	TC- - -GTGAAATGTAAGAGAAATTAATCTGAAATAAATTAATTTCTGTTAACTGGGTCA	2137
Qy	5379	TTCTCTCTCATCACTGAGCATCTGCTTTTTTGGAGGCCCTTGGGTAGTAGTGGAGATG	5438
Db	2138	TTTCTTCTCTATGCATGAGCATCTGCTCTGTGGAAGGCCCAAGGAATAGTAGTGAGATA	2197
Qy	5439	CTAAGGTAAAGCCAGACTCATACCCACCATAGGGTCGTAGAGCTTAGGAGCTGCAGTCCAC	5498

Db	2198	CTAGGCTAAGCCAGACACACACCTACCGATAGGGGTATTAAAGACTGTAGGACCGCGGTCAAT	2255
QY	5499	GTAATCAAGGTGGCAGAGATGTCTCTAAAGATGTAGGGAAAGTCAGAGAGGGGTGAGGG	5558
Db	2258	ATAATTAAGGTGACAAAGATGTCTCTAAGATGTAGGGGAAAAAGT- ---AACGAGTGTGGG	2313
QY	5559	TGTGGGGCTCCGGGTGAGAGTGGTGAGAGTGCATTCGCTCAGCTGAGTGGGGCATTTTGGGCT	5618
Db	2314	TATGGGGCTCCAGGTGAGAGTGGTCCGGGTGTAAATTCCTCTGTG-TGGGGCCTTTTGGGCT	2372
QY	5619	TTGGGAAACTGCAGATCTCCCTTCCTGGGGAGCTGATTGTTAATCATCTCTGGGTGGATCC	5674
Db	2373	TTGGGAAACTCCATTTTCTCTGAGGGATCGATTCTTAATGAAGCTGGTGGGTCC	2428
RESULT 14			
Q32357			
ID	Q32357	standard; cDNA; 2531 BP.	
AC	Q32357		
DT	22-APR-1993	(first entry)	
DE	MAGE-4 gene.		
KW	melanoma antigen; MAGE TRA; melanoma antigen tumor rejection antigen;		
KS	tumor rejection antigen precursor; MAGE; antigen E; gene family; ss.		
KW	Homo sapiens.		
EH	Key	Location/Qualifiers	
FT	cds	625..1578	
FT		/*tag= a	
PN	W09220356-A.		
PD	26-NOV-1992.		
PF	22-MAY-1992:	U04354.	
PR	23-MAY-1991:	US-705702.	
PR	09-JUL-1991:	US-728838.	
PR	23-SEP-1991:	US-764364.	
PR	12-DEC-1991:	US-807043.	
PR	(LUDW-) LUDWIG INST CANCER RES.		
PI	Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;		
PI	Van Den Eynde B, Van Der Bruggen P, Van Pel A;		
DR	WPI: 92-415460/50.		
PT	Nucleic acid mol. encoding a human tumour rejection antigen		
PT	precursor useful as an immunostimulant in a vaccine for		
PT	treating and preventing cancers, also useful in diagnosis		
PS	Disclosure: Page 79-80; 142pp: English.		
CC	The sequences given in Q32352-69 represent a new family of genes		
CC	referred to as melanoma antigens (MAGE). The cDNAs of this gene		
CC	family were identified during the isolation of the antigen E gene.		
CC	The MAGE cDNAs, when tested, did not transfer expression of antigen		
CC	E, but they did show substantial homology to the antigen E cDNA		
CC	sequence. The MAGE DNAs share a certain degree of homology with each		
CC	other and are expressed in tumour cells including several types of		
CC	human tumor cells as well as in human tumors. MAGE expression is not		
CC	restricted to melanomas. MAGE refers to a family of tumor rejection		
CC	antigen precursors. The antigens resulting from these genes are		
CC	referred to as MAGE TRAs or melanoma antigen tumor rejection antigens.		
CC	See also Q32351.		
SC	Sequence 2531 BP;	608 A; 592 C; 701 G; 630 T;	
Query Match			
Best Local Similarity 29.6%; Score 1677.2; DB 1; Length 2531;			
Matches 2073; Conservative 0; Mismatches 318; Indels 65; Gaps			
QY	3256	GGATCAGGCCCTGCCAGGAAAAATATAAGGSCCTCGGTGAGACAGAGGGGTGATCC	3315
Db	1	GGATCAGGCCCTGCCTGCAGAAATGTGAGGSCCTGAGTGAACACAGTGGGGATCATCC	60
QY	3316	ACTGCATGAGATGGGGATGTCCAGAGATCCAGGCCACCTCTCTGGTAGCAGTAGAGC	3375
Db	61	ACTCCATGAGATGGGGACCTCACAGATCCAGCCTACCTCTTGTATGGCACTGAGGAC	120
QY	3376	CAGGGCTGTGTGGGTTCGACACCTCGAGGCCCGGTGGATTCTCTCTCTGGAGCTCCA	3435
Db	121	CGGGCTGTGTCTACAGTCTGCACCTTAGGGCCCATGATTCCTCTCTCTAGAGACTCCA	180

Search completed: December 11, 1999, 06:29:38
Job time: 12965 sec

11/11/99 08:20:19 1999

THIS PAGE BLANK (USPTO)